DOCUMENT RESUME

ED 404 288	SP 033 927

TITLE	Texas Teacher Appraisal System. Evaluation Study.
INSTITUTION	Texas Education Agency, Austin. Div. of Curriculum Development.
PUB DATE	Apr 91
NOTE	186p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	Administrator Attitudes; *Career Ladders; *Contracts; Elementary Secondary Education; *Evaluation Criteria; Evaluation Methods; Evaluation Problems; Surveys; Teacher Attitudes; Teacher Effectiveness; *Teacher Evaluation; *Teaching (Occupation)
IDENTIFIERS	*System Evaluation; *Texas Teacher Appraisal System

ABSTRACT

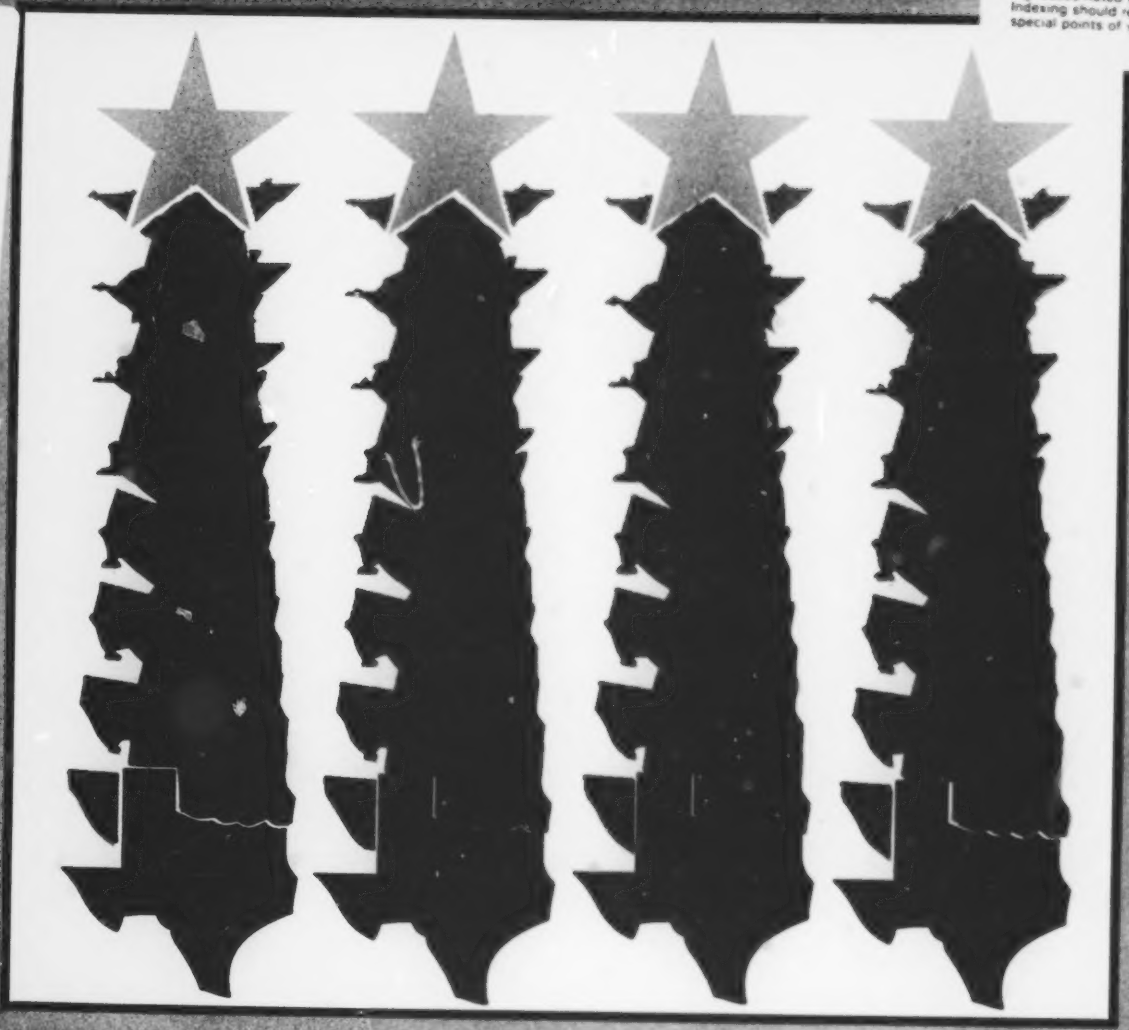
In 1984, consistent with reform efforts, the Texas legislature had directed the State Board of Education to adopt an appraisal process and criteria with which to appraise teachers' performance for career ladder purposes, improvement of instruction, and contract renewal decisions. In 1989, the Texas Legislature mandated an evaluation study of the Texas Teacher Appraisal System (TTAS) to be conducted by the Texas Education Agency. This report describes the implementation and results of the evaluation. The analyses generally indicated that the TTAS as implemented did not result in discrimination among various performance levels of teachers. This conclusion was based on the finding that use of the TTAS resulted in repetitive scores, patterning of score frequencies, and use of only two of the five available summary categories. Although most respondents agreed that the TTAS has potential as a teacher appraisal system, it is not being used in a manner consistent with good practice. There was also no support from the results of the study for keeping the career ladder as currently implemented in districts. Several approaches to improving the TTAS have potential, such as restricting the use of TTAS to beginning teachers, including objective measures of student performance, and making participation in the system voluntary. Contains several recommendations. (LL)

****
*   Reproductions supplied by EDRS are the best that can be made   *
*   from the original document.   *
****

ED 404 288

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: SP

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. EA



EVALUATION STUDY OF THE TEXAS TEACHER APPRAISAL SYSTEM

FROM THE STATE BOARD OF EDUCATION

TEXAS EDUCATION AGENCY
AUSTIN, TEXAS

APRIL 1991

SP033927

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
L Kemp
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

EVALUATION STUDY OF THE TEXAS TEACHER APPRAISAL SYSTEM

TEXAS EDUCATION AGENCY
DIVISION OF PROGRAM EVALUATION
AUSTIN, TEXAS

STATE BOARD OF EDUCATION
(State Board for Vocational Education)

CAROLYN HONEA CRAWFORD, Beaumont
Chairman of the State Board of Education
District 7

BOB AIKIN, Commerce
Vice Chairman of the State Board of Education
District 9

MARY HELEN BERLANGA, Corpus Christi
Secretary of the State Board of Education
District 2

Board Members

RAYMOND A. ALEXANDER, Houston
District 4

JACK CHRISTIE, Houston
District 6

EMMETT J. CONRAD, Dallas
District 13

WILL D. DAVIS, Austin
District 10

MONTE HASIE, Lubbock
District 15

WILLIAM L. HUDSON, Wichita Falls
District 14

GERALDINE MILLER, Dallas
District 12

JANE NELSON, Double Oak
District 11

RENE NUÑEZ, El Paso
District 1

MARY KNOTTS PERKINS, Lufkin
District 8

JOHN H. SHIELDS, San Antonio
District 5

ESTEBAN SOSA, San Antonio
District 3

THOMAS E. ANDERSON, JR.
Interim Commissioner of Education
(Executive Officer of the State Board of Education)

BEST COPY AVAILABLE

State Board Of Education
1701 North Congress Avenue
Austin, Texas 78701-1494
(512) 463-9007

Carolyn Honea Crawford, Ph.D.
Chairman
Beaumont, District 7

Bob Aikin
Vice Chairman
Commerce, District 9

Mary Helen Berlanga
Secretary
Corpus Christi, District 2

Raymond A. Alexander
Houston, District 4

Jack Christie, D.C.
Houston, District 6

Emmett J. Conrad, M.D.
Dallas, District 13

Will D. Davis
Austin, District 10

Monte Hasie
Lubbock, District 15

William L. Hudson
Wichita Falls, District 14

Geraldine Miller
Dallas, District 12

Jane Nelson
Double Oak, District 11

Rene Nuñez
El Paso, District 1

Mary Knotts Perkins
Tyler, District 8

John H. Shields
San Antonio, District 5

Esteban Sosa
San Antonio, District 3

Thomas E. Anderson, Jr., Ph.D.
Interim Commissioner of Education
(512) 463-8985

April 1991

The Honorable Ann W. Richards, Governor of Texas
The Honorable Bob Bullock, Lieutenant Governor of Texas
The Honorable Gibson D. Lewis, Speaker of the House
Members of the 72nd Legislature:

The Texas Education Agency has conducted an evaluation of the Texas Teacher Appraisal System (TTAS). This study was authorized by Senate Bill 417, Section 2.06.

The State Board of Education hereby submits the report. Additional copies are available from the Publications Division at the Texas Education Agency, 512/463-9744.

Respectfully submitted,

Carolyn Crawford

Carolyn Honea Crawford, Chairman
State Board of Education

BEST COPY AVAILABLE

ACKNOWLEDGEMENTS

This report was prepared by the Texas Education Agency, Division of Program Evaluation to summarize the results of its evaluation of the Texas Teacher Appraisal System.

The Texas Education Agency appreciates the efforts of the many teachers, principals, and superintendents who participated in the data collection for the study. In addition, thanks go to the representatives of the various professional organizations who contributed their time to interviews regarding this evaluation.

This study was conducted with considerable support from Lynn M. Moak, Deputy Commissioner for Research and Development and staff in the divisions of Teacher Education and Management Assistance and Personnel Development, working with Richard Swain, Assistant Commissioner for Professional Development. In particular, Gary Dill of the Division of Teacher Education provided valuable contributions to this evaluation.

Project Staff:

Division of Program Evaluation

David Stamman
Division Director

Cherry Kugle
Project Manager

Linda Hargrove
Lynn Trent
Martha Perez
Leigh Sharpe
Misti MacAfee

Recommendations for policy changes in this report were obtained during the course of this evaluation. Some of these ideas and recommended actions may not necessarily reflect the viewpoints of the Texas Education Agency or the State Board of Education.

BEST COPY AVAILABLE

CONTENTS

Page

EXECUTIVE SUMMARY 1

INTRODUCTION 7

RESULTS

CHAPTER 1: Conceptual Structure of the TTAS - The Texas Teacher Appraisal System as a System 15

CHAPTER 2: Implementation of the TTAS 25

CHAPTER 3: Subjectivity or Bias 33

CHAPTER 4: Uses of the TTAS 49

CHAPTER 5: Impact of the TTAS 55

CHAPTER 6: Career Ladder 63

OVERALL SUMMARY 69

APPENDICES

Appendix A: Historical Review and SBOE Rules 71

Appendix B: Contractor Review of Published Literature 93

Appendix C: Opinion Survey and Instructions 143

Appendix D: Interview Protocol and List of Districts 149

Appendix E: Analyze Categories for TTAS and Career Ladder 157

Appendix F: Recent Changes in TTAS and Comment from Assistant Commissioner for Professional Development 165

Appendix G: Independent Report from Contractor Concerning the Texas Teacher Appraisal System 171

LIST OF TABLES AND FIGURES

Page

FIGURE 1: Decision Points for Categorization of Response Patterns... 12

TABLE 1: Survey Response Rate.................................... 12

TABLE 2: Sample and Statewide Distribution of Teachers............. 13

TABLE 3: Deviation From Normal Curve.............................. 39

TABLE 4: Points in Total Score Distribution With Deviations From Pattern... 40

TABLE 5: Number and Percent of Teachers in Various Categories...... 41

TABLE 6: TTAS Mean Score For All Teachers.......................... 41

TABLE 7: Total TTAS Score by Teacher Demographics.................. 43

TABLE 8: Total TTAS Score by Experience and Degree................. 44

TABLE 9: TEAMS Scale Scores From 1985-1989......................... 58

TABLE 10: TEXAS SAT and ACT Scale Scores........................... 59

TABLE 11: Data From Selected Districts............................. 60

TABLE 12: Career Ladder Assignment by Teacher Demographics.......... 63

TABLE 13: Percent of Teachers on Career Ladder Steps (Wealth Categories)................................... 65

TABLE 14: Percent of Teachers on Career Ladder Steps (District Size and Type)............................... 65

EVALUATION OF THE TEXAS TEACHER APPRAISAL SYSTEM
Executive Summary

Senate Bill 417, passed by the 71st Texas Legislature (Regular Session) in the spring of 1989, mandated an evaluation study of the Texas Teacher Appraisal System (TTAS) to be conducted by the Texas Education Agency (Agency). In the reform legislation of 1984, the 69th Legislature directed the State Board of Education to adopt an appraisal process and criteria with which to appraise the performance of teachers for career ladder purposes. The impetus behind a mandated appraisal system was a desire to standardize appraisal of teachers in the state so that the newly mandated career ladder could be implemented in a more uniform manner than would be possible with the wide variety of appraisal approaches that were being used. In addition to career ladder assignments, information from these appraisals was to be used for improvement of instruction and for contract renewal decisions.

Because monetary resources for this evaluation were relatively low and Agency resources in terms of statewide data and data-analysis capability were relatively high, an evaluation approach was selected which would maximize the use of these resources and ensure an independent assessment. Thus, the evaluation was conducted as a cooperative effort between the Texas Education Agency and an external contractor.

Major components of the study included:

1. completion of a research study to assess the current literature;
2. development and administration of an opinion questionnaire to 2,000 teachers and administrators which generated 1,473 responses;
3. development and implementation of 37 interviews at ten sites;
4. quantitative and qualitative analysis of survey and interview data;
5. statistical analysis of data contained in Public Education Information Management System (PEIMS) computer files (approximately 180,000 teachers); and
6. application of the findings relative to current practice.

Five major themes were identified from the analyses of data and will be separately presented, followed by a brief discussion of information on the career ladder. Findings within each of these themes are thoroughly discussed in the main report and result in a number of recommendations concerning the TTAS. The recommendations are repeated here and serve as discussion points for future policy decisions.

Findings

Texas Teacher Appraisal System as a System. The TTAS was carefully designed to have the majority of features found in most accepted appraisal systems. Opinions indicated that the TTAS is suited for all grade levels and for assessment of basic teaching practices. There was considerably less support for assessing teacher performance with differing ability level students and in teaching critical thinking skills; in fact, teachers were negative regarding the first of these issues and neutral in opinion on the latter.

Even though there was support for the ability of the TTAS to discriminate polar extremes of effective and ineffective teaching practices, statewide PEIMS data indicate that this ability is not being used. Opinion was mixed concerning the use of student outcomes to assess effective teaching, with administrators favoring this practice and teachers opposing it.

Implementation of the TTAS. There was a very clear difference of opinion between administrators and teachers concerning the award of exceptional quality points (EQs). While both groups favor the continued award of EQs at the criterion level, teachers tend to support the idea of awarding EQs at the indicator level as well. Administrators were very much against this idea, possibly due to the requirement of documentation for each award. A potential advantage to awarding EQs at the indicator level would be an increase in the dispersion of scores; however, it is not clear this would occur. Simply raising the total possible score without a concomitant increase in variance among teacher appraisal scores will not add to the effectiveness of the TTAS.

There is a great deal of support for conferences between teachers and appraisers. Because an improvement plan must only be written when scores are less than satisfactory and almost none are, the conference becomes the primary means to formally transmit suggestions and comments. Since the lower categories are almost never used and the middle category is used rarely, the implication is that written, formalized expectations for improvement may not be a viable part of this system.

The majority of respondents believe that there are problems with either the TTAS or its implementation. By combining categories of responses, it appears that 64 percent of teachers and 48 percent of administrators believe that there are major problems with the TTAS. Additionally, 50 percent of teachers and 35 percent of administrators believe that there are problems with implementation. It might be tempting to say that there are fewer problems with implementation than with the system itself, but the percentages of respondents indicating problems are fairly high in either case.

Subjectivity and Bias. While the responses of administrators indicated that the TTAS does indeed reflect actual teaching ability and is fairer than previous systems, teachers did not share this perception. Respondents agreed that observable and quantifiable teaching behaviors are well-described and defined by the TTAS. There was less confidence that quality of teaching is measurable with this system. Respondents to the survey indicated that the TTAS reflects teaching and contains needed definitions, but that scores may not be objectively determined.

It was felt that monitoring and training of appraisers is adequate and that interpretations of classroom behavior are fairly consistent. Appraisers' pedagogical knowledge was not questioned, but there was less support that appraisers have appropriate subject area knowledge.

It is evident from examination of PEIMS data that deviations in patterns and repetition of scores were consistent with the premise that influences external to the instrument affect the scores. Even though an explanation is not proposed for this patterning, it is clear that the strength and regularity of the score patterns point to such external influences. The mean scores for teachers were quite high and, for all practical purposes, only two classifications were actually used - "exceeds expectations" and "clearly outstanding." A very large percentage of teachers received identical scores over four time periods with, it is assumed, different appraisers. This highly unlikely repetition of scores is an indicator of a problem with the system. The TTAS, as practiced, is over-reliable to the extent that its validity is highly questionable.

No matter how carefully an observation system is designed in terms of standardization and objectivity, there always exists the potential for appraiser bias. Two external forces which might influence scores on the TTAS are immediately apparent. The first is the inherent correspondence between TTAS scores and career ladder decisions. Given the need for high levels of performance on the TTAS in order for teachers to move up the career ladder, there exists a pressure for upward bias in the scores. Appraisers are likely to be reluctant to produce unfavorable TTAS ratings which would then preclude career ladder bonuses to teachers who, it is generally agreed, are underpaid. The second force driving TTAS scores upward is the reluctance to assign a person to the category "meets expectations" since this is equivalent to average performance. Although most human behavior is, by definition, average, until it can be accepted that "average" or "meets expectations" is an acceptable state of affairs, there will be pressure for higher scores.

Uses of the TTAS. There was mixed opinion concerning the association of the TTAS with the career ladder. It is much more evident that the use of a different scoring procedure was not supported and the addition of other domains was questioned as well. There was moderate support for the use of a different instrument and procedure. There was no support for campus assignment to be made based on results from the TTAS.

Perhaps the question of most interest is the one concerning contract renewal and termination. Principals are moderately in support of this use of the TTAS, teachers are moderately against it, and the superintendents are in the middle on this issue. According to Webster (1988), "negligent evaluation" is one of the ways in which employees use evaluations to sue their employers. As is evident from earlier data, the majority of teachers are rated in the highest categories, leading to little ability to discriminate among teachers based on TTAS scores. Given the apparent evidence that the TTAS system, as practiced and regardless of concept or design, is influenced by factors external to the system, the denial of contracts or merit pay on the basis of TTAS scores may be vulnerable to litigation.

Impact of the TTAS. Although administrators agree that improvement in teaching skills can be attributed to the TTAS, only half of the teachers agree that this is so. There is less support for the question whether the TTAS has improved unsatisfactory teaching skills. There is no support from the survey results that student achievement has been affected by the TTAS and little support from student performance data that the TTAS has been instrumental in creating gains. From the available data it is clear that if the TTAS was meant to increase teacher effectiveness which would in turn increase student performance, this goal was not met.

Summary of Recommendations

The following major recommendations are keyed to the results chapters within the report and are derived from published research, data from the survey, from on-site visits and from PEIMS. These recommendations are discussed in the body of the report which should be carefully reviewed before actions based on these recommendations are undertaken.

Recommendation 1.1.	Investigate the relationship of student outcomes, especially as included in the Academic Excellence Indicator system, to individual teachers and to campus-level aggregations of appraisal scores.
Recommendation 1.2.	Investigate the potential for adding new domains or criteria to the TTAS which would assess the teachers' abilities to teach critical thinking skills and to teach students of different ability levels.
Recommendation 2.1.	Establish a pilot program in selected districts under careful controls to examine awarding EQs at the indicator level.
Recommendation 2.2.	Establish a pilot program in selected districts under careful controls to examine documentation of EQs. Restriction, enhancement, or removal of requirements for documentation should be examined.
Recommendation 2.3.	After the first two years of teaching, eliminate annual evaluations. Appraise every third year for a minimum of four observation periods, only one of which would be scheduled.
Recommendation 3.1.	Disassociate the TTAS from the career ladder.
Recommendation 3.2.	Evaluate appraisers based on a distribution of scores awarded over time, especially as compared to statewide information. Publish means and distributions of campus, district and statewide TTAS scores.
Recommendation 3.3.	Continue the educational process for appraisers, teachers, and the general public concerning the meaning and utility of the various rating categories. Universities associated with teacher training need to strengthen curriculum in this area.
Recommendation 3.4.	Investigate the use of a cadre of "professional" appraisers, perhaps based at the education service centers.
Recommendation 4.2.	Do not use TTAS scores for campus assignments.
Recommendation 4.3.	Do not use the TTAS alone to make contract decisions.
Recommendation 4.4.	Create a voluntary appraisal program linked to incentive pay.
Recommendation 5.1.	Create two comparison groups of districts and campuses by continuing the TTAS in its present form in one group and replacing it with

"walk-through" visits on a regular basis in the other group. Examine score trends for overall student performance on a variety of measures.

Recommendation 5.2. If changes are made to the TTAS or the career ladder, then a reassessment similar to this evaluation should be repeated at some point in time after implementation of the changes.

Career Ladder

Results of a question which was included in the TTAS questionnaire:

OPINION PERCENTAGES WITHIN PROFESSIONAL CATEGORY

Options for the Career Ladder	Overall	Principals	Superintendents	Teachers
Abolish; Provide Monetary Incentives to High-Performing Campuses	11.2%	12.7%	14.1%	9.4%
Abolish	18.5%	22.3%	27.9%	12.3%
Abolish;Increase Salaries	41.8%	44.9%	44.9%	36.7%
Retain, as is	2.3%	1.9%	1.1%	3.1%
Retain, and Fully Fund	21.3%	14.6%	8.0%	31.9%
None of the Listed Options	5.0%	3.6%	4.0%	6.8%

Interestingly, opinions did not vary by the level of the career ladder to which the teacher was assigned nor by the years of experience of the educator. Support for maintaining the career ladder as it currently exists is minimal at best. Administrators clearly favor abolishing the career ladder with about one quarter indicating it should go, even with no alternative. Not surprisingly, almost half of the administrators and slightly over one-third of the teachers want to see the career ladder abolished and replaced with increased salaries. It was somewhat surprising to see that almost one-third of the teachers wanted to retain the ladder, but to fully fund it, an option not espoused by the administrators.

Recommendation: Given a certain limit on funding, a compromise position which considers the results of the survey would be to abolish the career ladder, and determine alternative methods of allocating funds.

EVALUATION OF THE TEXAS TEACHER APPRAISAL SYSTEM
Introduction

Senate Bill 417, passed by the 71st Texas Legislature (Regular Session) in the spring of 1989, mandated an evaluation study of the Texas Teacher Appraisal System (TTAS) to be conducted by the Texas Education Agency (Agency). This report describes the implementation and results of the evaluation.

History of the Texas Teacher Appraisal System

In the reform legislation of 1984, the 69th Legislature directed the State Board of Education to adopt an appraisal process and criteria with which to appraise the performance of teachers for career ladder purposes. The impetus behind a mandated appraisal system was a desire to standardize appraisal of teachers in the state so that the newly mandated career ladder could be implemented in a more uniform manner than would be possible with the wide variety of appraisal approaches that were being used. In addition to career ladder assignments, information from these appraisals was to be used for improvement of instruction and for contract renewal decisions, especially for beginning teachers. After a timeline was adopted in 1985, an advisory group met to advise the Agency on issues concerning evaluation, measurement, and procedures. Agency staff conducted an extensive review of the literature on teaching effectiveness and surveyed the appraisal systems in use at that time in other states and in a number of districts in Texas. Following a job-relatedness survey, pilot studies were conducted during the 1985-86 school year in six school districts. The State Board of Education adopted rules governing the teacher appraisal system and, in late spring of 1986, training of state appraisal trainers commenced. By the end of summer 1986, 13,000 appraisers had received training and, in the fall, orientations for teachers were conducted. During the winter and spring of 1987, standards were adopted and refinements were made to the appraisal system. Further refinements continued for the next two years in the areas of performance standards, scoring procedures, observation methods, and grievance procedures. Appendix A provides a more detailed historical overview of the development of the TTAS, the assumptions underlying the system, and the statutory citation and board rule pertaining to the TTAS. In addition, an overview of the TTAS as implemented during the 1989-90 school year is provided.

Components of the Study

Because monetary resources for this evaluation were relatively low, but resources at the Agency were relatively high in terms of statewide data and data-analysis capability, an evaluation approach was selected which would maximize the use of the resources available. Thus, the evaluation was conducted as a cooperative effort between the Texas Education Agency and a contractor selected through the standardized Request For Proposal (RFP) process. After the RFP was issued in May, 1990, and proposals were reviewed by agency staff in July, 1990, MGT of America, Inc. was selected as the contractor for this evaluation.

BEST COPY AVAILABLE

The evaluation of the current TTAS consisted of several components, some of which were the primary responsibility of the contractor and some of which were completed by agency staff in the Division of Program Evaluation. The contractor participated at some level in all phases of this study. The components of the study are as follows, with the responsible party indicated in parentheses:

1. completion of a research study to assess the currently available literature and the systems in place or proposed in other locales (contractor);

2. development (contractor) and administration (agency) of an opinion questionnaire to teachers and administrators concerning the implementation and uses of the TTAS ;

3. development and implementation of on-site interviews with a sample of teachers and administrators (contractor);

4. quantitative and qualitative analysis of the data obtained from the mailed questionnaires (agency) and personal interviews (contractor);

5. statistical analysis of data contained in computer files to evaluate district TTAS and career ladder information (agency);

6. application of the findings relative to current practice in order to examine differences between ideal and actual standardized appraisal systems (agency and contractor).

Component 1: Review of the Literature

The contractor was instructed to study all existing literature concerning appraisal systems in education in order to assess appraisal systems in place or proposed in other locales. In addition, the use of appraisal systems in other industries was reviewed for pertinence to educational systems. Detailed information concerning current practice and proposed systems in other states and local districts is provided. The goal of this component was to provide a basis from which the TTAS could be evaluated relative to other systems. The contractor was solely responsible for this component of the study. Completion of the first component of the study resulted in a comprehensive review of the literature which is presented in Appendix B.

Components 2 - 5: Methodology

Opinion Questionnaire. Development of the opinion questionnaire for this evaluation began with interviews by the contractor of individuals knowledgeable about the TTAS, many of whom represented educator groups which included the following: Association of Texas Professional Educators, Texas Association of Secondary School Principals, Texas Classroom Teachers Association, Texas Elementary Principals and

Supervisors Association, Texas Federation of Teachers, and Texas State Teachers Association. Through this process, a number of major issues emerged which were then used by the contractor, working with agency staff, to develop the themes, and ultimately the items, used in the opinion questionnaire. The themes identified were the following:

- the impact of the TTAS on teachers and teaching practices, i.e., whether the TTAS makes a difference in quality of teaching;
- the TTAS instrument content and indicators;
- the TTAS scoring procedures and measurement scale used;
- the TTAS processes of observation and appraisal;
- the qualifications of the TTAS appraisers.

Initially, a three-point scale of "extensively," "somewhat," and "not at all" was chosen as the response format. However, in order to provide a choice between "extensively" and "somewhat," and to produce a scale which would allow placement of responses into positive and negative groups, the final response format consisted of "completely," "adequately," "inadequately," and "not at all." In addition to the forced-choice questions, open-ended responses were solicited by providing an opportunity at the end of the questionnaire for the respondents to make general comments. The final instrument and instructions for its completion are included as Appendix C.

A sample of 2000 teachers and administrators employed during the 1989-90 school year was selected from the Public Education Information Management System (PEIMS) to receive the opinion questionnaire. The sample was composed of 1,000 teachers with appraisal scores drawn at random from the PEIMS database, 337 superintendents from the districts where the teachers were employed, and 663 principals from the same districts. One hundred eighty-four of the principals were from the same campuses as the teachers. Questionnaires were mailed to the superintendents, with instructions to distribute them to the educators on an enclosed list. Each educator received a return envelope addressed to the Division of Program Evaluation at the agency so that responses would be confidential. No identifying information was requested from the respondents except the general demographic data shown in questions one through seven on the form. District demographic information was coded onto each form so that responses could be analyzed by general categories such as region of the state (five groups), average daily attendance (six groups), percent of minority students (three groups), and wealth of district (three groups).

Interviews. An interview protocol was developed by the contractor which followed the same themes as those in the opinion questionnaire. The purpose of the interviews was to solicit additional comments to clarify opinions regarding the uses and implementation of the TTAS. In addition, the interviews allowed the exploration of issues related to the TTAS which were not specifically addressed in the questionnaire. From the sample of 337 districts included in the opinion survey, ten districts were selected to serve as interview sites which were representative of the diversity of districts in the state. Geographical location and size were the primary determinants in this selection. At each district, individual confidential interviews were conducted with the superintendent, an appraiser, and two teachers; a total of 37 educators were interviewed. A copy of the interview guide and the list of districts where interviews were conducted is provided in Appendix D.

Data Analysis. Both statistical and qualitative assessments of the data obtained from the mailed questionnaires and personal interviews were made. Staff from the agency conducted the analyses of the questionnaire data; the contractor analyzed the interview data. Quantitative analyses from the questionnaire consisted primarily of frequency tables constructed to allow an examination of responses grouped by such demographic information as sex, ethnicity, and position (i.e., teacher, principal, superintendent). Qualitative analyses from the questionnaire were performed by grouping the written comments solicited at the end of the instrument into the major themes which evolved from examination of the quantitative data.

In analyzing the interview data, the contractor calculated the percentages of "yes" and "no" responses to the interview questions, and summarized the numbers and percentages of types of comments received from the interviewees.

During a meeting between agency and contractor staff, patterns of responses from these data sources were examined and five major issues were identified as recurrent themes in the various data sources. These overlap with, but are not identical to, the themes identified during the developmental work of this evaluation. The TTAS issues which emerged from the data collection are the following:

- Conceptual structure of the TTAS
- Mechanics of implementation of the TTAS
- Subjectivity or bias in the TTAS
- Uses of TTAS results for other purposes
- Impact of the TTAS

To facilitate discussion of the results of this evaluation, responses from the questionnaire items, and qualitative responses from the questionnaires and interviews were grouped according to these five themes. Each chapter in the results section addresses one of these themes.

Analysis of District TTAS Data. Data from the Public Education Information Management System (PEIMS) were analyzed to examine statewide trends in TTAS scores over four measurement periods (Fall, 1988; Spring, 1989; Fall, 1989; Spring, 1990). In addition, data were obtained from selected districts through two education service centers in order to examine the relationships between teacher and student demographics, and teacher appraisals and student performance. However, the data obtained from the sample of districts proved to be too non-representative to allow confidence in the results and will not be reported here.

Appraisal scores and career ladder placements were extracted from the PEIMS data base and examined by several teacher characteristics, such as sex, ethnicity, and years of experience. In addition, scores were grouped by district characteristics using the ANALYZE program to examine patterns of scores across such district demographics as size, wealth, region, type, and student characteristics.

Components 2 - 5: Results

In this section of the report, analyses of the data will be presented in five chapters which correspond to the major issues identified in the course of the study. Although the five TTAS issues are clearly interrelated, they are separated and presented in distinct chapters to structure the reporting of a relatively large amount of information. Each of the chapters will present-information received from the structured questions in the survey, the open-ended responses from the survey combined with the results of on-site interviews, and information available from the agency's Public Education Information Management System (PEIMS). Preliminary recommendations which arise from these sources of information and opinion regarding the issues will be presented at the end of each chapter. A brief sixth chapter is included to provide information obtained concerning the career ladder. Following this results section, an overall summary will be provided. As mentioned previously, the five issues which emerged from the evaluation are the following:

- Conceptual structure of the TTAS
- Mechanics of implementation of the TTAS
- Subjectivity or bias in the TTAS
- Use of TTAS results for other purposes
- Impact of the TTAS

Within each of the five chapters, a common format will be used to present the questions from the survey and other information. Some questions from the survey may be paraphrased to simplify the presentation. In most cases, the responses "completely" and "adequately" will be combined as will "inadequately" and "not at all" to better present differences of opinion. Where a significant number of respondents indicated opinions at either extreme, this will be so noted in the discussion. The typical presentation of results from the survey will consist of opinions collapsed across response categories for each professional position: principal, superintendent, and teacher. Statistical information will be provided indicating significance of relationships among the various professional categories and the response patterns. The reader is cautioned that the large number of respondents will result in many statistically significant relationships. Thus, a more practical means of identifying significant findings was devised, using specific parameters selected to nominally rate the opinion patterns. Operationally, the parameters shown in Figure 1 were used to label distributions of opinion.

FIGURE 1
Decision Points for Categorization of Response Patterns

Percentage of respondents expressing a belief
Affect 100 70 60 50 40 30 0
Label : Strong :Moderate | Neutral |Moderate : Strong :
SUPPORT <--> NON-SUPPORT

For example, if 80 percent of the teachers expressed support, this was labeled as strong support for that position because a large portion, but not all, indicated "completely" or "adequately." Note that this describes the overall position of a group, not of particular individuals who may have very strong, and opposite, opinions. The classification of neutral may hide the fact that half are supportive, while half are not. No decision or recommendation can be supported given this type of opinion dichotomy. The neutral category comprises twenty percent of the overall range as does the moderate category. The categories labeled strong are reserved for the 30 percent at either end of the scale in an attempt to compensate somewhat for the central tendencies inherent in most opinion surveys. To be consistent, the delimiters between categories will be treated as absolutes; that is, 60.1 percent will be treated as moderate even though it is close to neutral.

Open-ended responses selected for inclusion in the results section are those which provide examples of the statements or positions asserted by those choosing to make a response. Because not all respondents included general comments, these must be considered additional rather than representative information.

Response Rate. The return rate for the surveys was reasonable with the lowest percentage received from teachers, two-thirds of whom returned completed questionnaires. Numbers and percentages of surveys returned are provided in Table 1.

TABLE 1
Survey Response Rate

Position	Number	Return Rate
Principal	488	73.6%
Superintendent	277	82.2%
Teachers	674	67.5%
Other	34	

For the total group which returned the surveys, the range of years of teaching or administrative experience was 1 year to 40 years with a similar range for local district experience. The most frequent response was six years of total experience and five years of local experience. Over all respondents, approximately 53 percent were female, 84 percent were white, ten percent were Hispanic and five percent were black.

For teachers, Table 2 presents the percent of the sample by sex and ethnicity compared to statewide distributions.

TABLE 2
Sample and Statewide Distribution of Teachers

SEX				ETHNICITY*					
State		Sample		State			Sample		
Male	Female	Male	Female	Black	Hisp	White	Black	Hisp	White
21.3	78.7	21.1	78.9	9.2	13.1	77.3	5.7	8.1	85.5

* Hisp - Hispanic. Other ethnic categories represent 0.4 percent within the state and 0.7 percent of the sample.

Percentages may not add to 100% due to rounding

The statewide data are for 178,427 teachers with appraisal scores in Spring of 1990; the sample data represent the 674 teachers who returned surveys. There is very good agreement between the state and the sample for males and females. White teachers are over-represented in the sample, but the absolute number of teachers in each category would argue for sufficient numbers of teachers to provide stability of the results. As will be discussed in the report, opinions tended to be quite stable across these demographic categories.

Note to the Reader: When examining the data presented in this section of the report, you will often see large differences of opinion between teachers and administrators. Recall that these differences will almost always be statistically significant due to the large number of respondents. However, from a qualitative point of view these differences are probably best explained by remembering that the TTAS is, somewhat simplistically put, something done by administrators to teachers. The perspectives of evaluators and those being evaluated will certainly differ, regardless of the type of evaluation system being used. In the case of the TTAS where performance is being evaluated in order to make decisions which affect bonus pay, contract renewal, and staff development activities, it should not be surprising to find somewhat less than enthusiastic attitudes toward the TTAS on the part of teachers. Thus, a somewhat constant difference between teachers and administrators might be expected. The use of the categorization approach described previously to assign labels to the opinion patterns should help determine when such differences are of practical significance.

Results

Chapter 1: Conceptual Structure of the TTAS

The Texas Teacher Appraisal System as a System

This part of the evaluation investigated this question: "Was the TTAS conceived in a manner which would lead to better appraisal of teacher performance?" It could be argued that all of the questions in the survey are pertinent to this issue; however, the items discussed in this section were selected as those that most directly address this question. The first set of four questions were asked from a perspective of "in your district." That is, these questions were to be answered based on the congruence of the TTAS to teacher appraisal in the respondent's own district. An additional set of questions to be examined in this chapter makes queries that are more general or conceptual in nature. The use of these two different perspectives should help to ascertain the perceptions of education professionals regarding the basic tenets of the TTAS as applicable to teachers and teacher appraisals.

Capabilities of the System

Survey Questions:

The complete questionnaire is included in Appendix C. Question numbers given here reflect those on the survey. For example, Q8(a) corresponds to question number 8, part a on the questionnaire.

Q8(a) To what extent does the TTAS reflect actual teaching ability or competence?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	57.7%	72.2%	59.8%	46.6%
Inadequately, Not at All	41.5%	27.8%	40.2%	53.4%
No Opinion	.8%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1428

Q8(e) To what extent does the TTAS identify ineffective teaching practices?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	63.9%	77.6%	65.6%	54.2%
Inadequately, Not at All	34.9%	22.4%	34.4%	45.8%
No Opinion	1.2%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1421

BEST COPY AVAILABLE

Q8(f) To what extent does the TTAS identify exceptional quality of teaching methods?

	Overall	Principals	Superin-tendents	Teachers
Completely or Adequately	60.6%	71.3%	58.8%	54.5%
Inadequately, Not at All	38.2%	28.7%	41.2%	45.5%
No Opinion	1.2%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1422

Q8(j) To what extent does the TTAS standardize appraisal of teachers?

	Overall	Principals	Superin-tendents	Teachers
Completely or Adequately	56.0%	67.6%	71.3%	43.3%
Inadequately, Not at All	41.7%	32.4%	28.7%	56.7%
No Opinion	2.3%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1406

When examining the results collapsed over all of the professional positions, and using the classification strategy that when 40 to 60 percent of respondents hold one opinion the affective significance of the responses is neutral, it can be seen that for items Q8(a) and Q8(j) approximately half supported these statements while approximately half did not. This would be classified as neutral when considering opinions as a whole. Overall, respondents were moderately supportive of items Q8(e) and Q8(f). Generally, overall opinions are not examined in detail in this report except to note where a large percentage declined to respond.

Principals strongly agreed with the capability of the TTAS to reflect actual teaching ability, identify ineffective teaching practices, and identify exceptional quality of teaching methods. Teachers were neutral in response to these questions while superintendents were moderately supportive of the ability of the TTAS to identify ineffective teaching practices. Interestingly, superintendents, who are less likely to be directly involved in the appraisal process, were very supportive of the statement that the TTAS standardizes appraisal while principals were moderately supportive and teachers neutral. It appears that the TTAS is considered to be appropriately constructed to reflect actual teaching ability, identify ineffective teaching practices and exceptional quality of teaching methods, but that these capabilities are not considered sufficient to standardize the appraisal process.

On-Site:

The on-site interviews conducted by the contractor allowed the respondents to comment upon the same questions as contained in the questionnaire, but with an opportunity to expand and clarify. In general, most of the interview results support the data obtained from the questionnaire. For the first three of the above questions, the responses collected from the on-site interviews did not differ significantly from those found in the questionnaire. There was moderately more support from interviewees for the notion that the TTAS does standardize the appraisal of teachers (74 percent said "yes," compared to an overall 56 percent agreement from the survey).

Open-Ended Comments:

(Superintendent) Use TTAS as a basic evaluation tool in all districts to establish a level of basic skills only.

(Superintendent) The TTAS does not lend itself to rate or capture the personal qualities that are so important to identifying an outstanding teacher.

(Superintendent) Range of scores for "meets" expectations" is too broad and should be narrowed from the bottom up.

(Teacher) More weight should be given to staying current in teaching trends, concepts, and methods.

(Teacher) The TTAS instrument is only as valid as the appraiser's knowledge and expertise.

PEIMS Data:

As will be discussed in depth later, data from PEIMS indicated that only 37 teachers out of almost 180,000 were identified as unsatisfactory and less than 0.2 percent were classified as below expectations. A very large percentage of teachers were rated as clearly outstanding, which tends to bring into question the ability of current practice to defensibly identify either ineffective or exceptional quality teaching. PEIMS data also indicated that there is a remarkable amount of stability in scores from one time to another. Unfortunately, this may not reflect reliability so much as predetermination of scores. Results of analysis of the PEIMS data indicate that the TTAS may over-standardize the appraisal process from a stability point of view, and may not achieve validity in terms of comparing observed teaching behaviors with standards of effective teaching.

Specific Attributes of the TTAS

Criteria and Indicators. On a more concrete level, questions were asked concerning the appropriateness of specific criteria and indicators to assess different instructional situations.

Survey Questions:

Q9(a) To what extent does the TTAS contain criteria and indicators which assess different ability levels of students?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	43.3%	50.8%	49.6%	36.1%
Inadequately, Not at All	54.5%	49.2%	50.4%	63.9%
No Opinion	2.2%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1407

Q9(b) To what extent does the TTAS contain criteria and indicators which assess basic teaching practices for most subjects?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	76.4%	83.7%	81.8%	70.8%
Inadequately, Not at All	22.0%	16.3%	18.2%	29.2%
No Opinion	1.6%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1415

Q9(c) To what extent does the TTAS contain criteria and indicators which assess teaching of critical thinking skills for students?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	51.9%	55.2%	47.4%	53.1%
Inadequately, Not at All	46.0%	44.8%	52.6%	46.9%
No Opinion	2.0%	n/a	n/a	n/a

By Position: Statistically Non-Significant; Chi-square F<.118 df=2 n=1409

Q9(c) To what extent does the TTAS contain criteria and indicators which assess different grade levels?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	61.1%	71.7%	67.5%	58.4%
Inadequately, Not at All	32.5%	28.3%	32.5%	41.6%
No Opinion	6.3%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1346

Administrators were neutral in their opinions concerning the ability of the TTAS to assess teaching activities as applied to different ability level students. Teachers, on the other hand, were moderately non-supportive of this feature. There was strong agreement from all respondents that the TTAS is adequate for basic teaching practices for most subjects. Neutral opinions were received from all types of respondents regarding the assessment of teaching critical thinking skills, while there was generally moderate and strong support for the use of the TTAS at different grade levels.

On-Site:

Interviewees expressed agreement regarding the ability of the TTAS to assess different ability levels of students and there was overwhelming agreement (95 percent) that the TTAS does assess basic teaching practices. Results from the remaining two questions in this section continued the trend for this group of respondents to be more positive in their view of the TTAS with 12 to 20 percent more of the interviewees, relative to survey respondents, indicating agreement that the TTAS has the ability to assess teaching behaviors for critical thinking skills and for various grade levels.

Open-Ended Comments:

(Principal) The TTAS should be changed to better fit a variety of teaching models and subject areas.

(Teacher) Fewer indicators - more description of what I do or need to be doing.

Teacher Specialization

Specialized Subject Areas. Four questions were included that asked whether certain specialized subject areas were adequately covered, i.e., is the assessment of teachers with responsibilities in these areas within the capability of the TTAS? This question produced no opinion from 30 to 40 percent of the respondents. Of those who did indicate an opinion, both on the questionnaire and in on-site interviews, there was an even split of opinion, regardless of position. The support for the ability of the TTAS to adequately assess the areas of art, vocational education, laboratories, and special education was neutral, at best.

Descriptions and Definitions

Clarity of Description. Another critical aspect of the TTAS is the clarity of descriptions and definitions of various teaching practices and behaviors. A system which does not clearly define the behaviors to be observed and rated cannot be reliable. The next five questions address this issue.

Survey Questions:

Q12(a) To what extent does the TTAS clearly describe or define the criteria for appraising effective teaching practices?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	77.9%	86.2%	83.6%	70.9%
Inadequately, Not at All	21.0%	13.8%	16.4%	29.1%
No Opinion	1.1%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1422

Q12(b) To what extent does the TTAS clearly describe or define the indicators for appraising effective teaching practices?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	77.2%	85.5%	83.6%	70.2%
Inadequately, Not at All	21.4%	14.5%	16.4%	29.8%
No Opinion	1.4%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1419

Q12(c) To what extent does the TTAS clearly describe or define observable teaching behaviors?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	77.3%	86.8%	84.7%	68.4%
Inadequately, Not at All	21.7%	13.2%	15.3%	31.6%
No Opinion	1.0%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1424

Q12(d) To what extent does the TTAS clearly describe or define quantifiable teaching behaviors?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	64.3%	76.1%	65.9%	57.1%
Inadequately, Not at All	33.4%	23.9%	34.1%	42.9%
No Opinion	2.3%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1406

Q12(e) To what extent does the TTAS clearly describe or define the quality of teaching performance?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	54.2%	58.4%	50.9%	47.3%
Inadequately, Not at All	38.0%	41.6%	49.1%	52.7%
No Opinion	7.8%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1422

There seems to be little doubt that the descriptions and definitions contained within the TTAS are perceived by educators to be very clear. In close correspondence with these results was the strong support from the respondents to Question 12(c) that the TTAS clearly describes or defines observable teaching behaviors. The question concerning quantifiable teaching behaviors evoked strong support from principals, moderate support from superintendents and more neutral support from teachers. The picture is not so clear when asked about the quality of teacher performance. In this case all three groups indicated support in the neutral range.

On-Site and Open-Ended Comments:

The information from the on-site interviews closely parallels the responses from the questionnaires. No significant comments were noted from the open-ended comments which yielded additional information on this topic.

Classification of Teachers

Discrimination Power of the TTAS. Two questions were included which explored the ability of the TTAS to discriminate between good and bad teachers and to assess the full range of teaching practices.

Survey Questions:

Q15(l) Within the TTAS, is there an ability for appraisers to discriminate polar extremes of effective and ineffective teaching practices?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	71.6%	81.8%	74.0%	68.2%
Inadequately, Not at All	24.7%	18.2%	26.0%	31.8%
No Opinion	3.7%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1384

Q15(m) Within the TTAS, is there an ability for appraisers to discriminate the full range of teaching practices?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	59.2%	69.4%	59.1%	56.2%
Inadequately, Not at All	37.3%	30.6%	40.9%	43.8%
No Opinion	3.4%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1388

BEST COPY AVAILABLE

On-site and Open-Ended Comments:

These sources of information tended to confirm the same pattern of findings as seen in results from the survey.

PEIMS Data:

If there is an ability of the TTAS to discriminate the polar extremes of teaching practices, analyses from the PEIMS data show that it is not being used. As will be discussed in detail later, there were practically no teachers rated in the lowest two categories on the TTAS and a predominance of teachers existed in the highest two categories.

Performance as Measured by Student Outcomes

Student Outcomes. Application of the TTAS consists of observations of teacher behaviors, termed the process of instruction. That is, assessment through observations of teacher behaviors which have been associated with higher student outcomes is the mechanism for making judgements concerning teacher effectiveness. Obviously in the classroom there are behaviors other than the teacher's. Observation of student behaviors is included in current appraiser training, but there exists no objective measurement of student outcome performance within the TTAS. One survey question explored opinions concerning the inclusion of student outcomes in the system. This would involve the addition of an objective measure to a subjective, albeit structured, system.

Survey Questions:

Q17 To what extent should student outcomes be included in the appraisal of teachers?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	48.8%	66.4%	65.2%	39.1%
Inadequately, Not at All	42.4%	33.6%	34.8%	60.9%
No Opinion	8.8%	n/a	n/a	n/a

By Position: Statistically *Significant*; Chi-square F<.001 df=2 n=1314

Differences in opinion between administrators and teachers were again apparent with moderate support for this proposition given by principals and superintendents and moderate support against this proposition evidenced by teachers. Administrators may feel that the inclusion of objective data would reduce some of the judgmental pressure inherent in a classroom observation system, while teachers may be concerned about the accuracy of measurement of student performance or about the reasonableness of attributing student performance, with its multivariate causes, solely to teachers.

On-site and Open-Ended Comments.

Again, results from these sources were supportive of the findings from the survey.

Another set of questions was asked concerning the TTAS and its implementation, but these will be covered in a later section concerning the mechanics of implementation of the TTAS.

Discussion and Recommendations

The review of the literature in Appendix B describes in detail the features of a credible appraisal system. The TTAS was carefully designed with most of these features in accordance with existing literature. There was reasonable support by administrators that the TTAS has these desired characteristics. Teachers were somewhat more neutral in their opinions. The opinions of respondents indicated that the TTAS is suited for all grade levels and for assessment of basic teaching practices. There was considerably less support for teaching different ability levels of students and critical thinking skills to be assessed by the TTAS; in fact, teachers were negative regarding the first of these factors and neutral in opinion on the latter. There was very strong support for the clarity of descriptions and definitions within the TTAS.

Although there was support for the ability of the TTAS to discriminate polar extremes of effective and ineffective teaching practices, and assuming that classroom observations should reveal an array of teacher practices, statewide appraisal scores from the PEIMS database indicate that this ability has not been effectively used. Opinion was mixed concerning the use of student outcomes in the appraisal of teachers with administrators favoring this practice and teachers opposing it.

The literature regarding this topic indicates that evaluation systems should be designed to limit the cognitive demands placed on appraisers and to match the requirements for accurate appraisal to the capabilities of the appraisers (Gaugler & Thornton, 1989). Capabilities in this case are those such as memory capacity, intelligence, and orientation to detail (Heneman, Wexely & Moore, 1987). Limiting the number of dimensions appraisers are required to process (Gaugler & Thornton, 1989), limiting the number of ratings per dimension (Lupton, 1989; Fox 1987/88), assuring familiarity with the job to be appraised (Hahn & Dipboye, 1988; Heneman, Wexely & Moore, 1987), and establishing a frame-of-reference for evaluation (Daley, 1987) have been shown to improve accuracy. Further, Brandt (1990) recommends the use of multiple criteria, multiple judges, and multiple data sources for teacher appraisal, including the observation of a variety of situations, the objective and systematic collection of representative samples of relevant data from a variety of sources, and student performance data.

Recommendation 1.1. Investigate the relationship of student outcomes, especially as included in the Academic Excellence Indicator System, to individual teachers and to campus-level aggregations of appraisal scores.

Although the relationship of student outcomes to individual teacher behaviors has always been problematic because of the potential instability of outcome measures and the differences in student abilities, there is a need to carefully assess the potential of this area of exploration.

Recommendation 1.2. Investigate the potential for adding new domains or criteria to the TTAS which would assess the teachers' abilities to teach critical thinking skills and to teach students of different ability levels.

The recommendation would require extensive input from universities and cooperation from districts to investigate this issue which is complicated in construct and in measurement.

Results

Chapter 2: Implementation of the TTAS

Any system of teacher appraisal, regardless of how well conceived or designed, must be implemented in a real world educational setting in the district. The accuracy and usefulness of the TTAS depends upon an understanding and acceptance, by the appraiser and the teacher, of the basic tenets of the system as used in the classroom. The questions reported in this section attempt to investigate the mechanics of implementation.

Higher Levels of Performance

Exceptional Quality Points. The TTAS allows the awarding of exceptional quality points (EQs) at the criterion level if specific teaching behaviors are demonstrated during the classroom observation. Three questions probed the use of the EQs.

Survey Questions:

Q14(a) To what extent is it appropriate to award exceptional quality points for each indicator in a criterion?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	38.5%	25.3%	28.7%	59.2%
Inadequately, Not at All	55.6%	74.7%	71.3%	40.8%
No Opinion	5.9%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1353

Q14(b) To what extent is it appropriate to award exceptional quality points for the entire criterion in a holistic way (current practice)?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	69.5%	81.4%	75.2%	65.2%
Inadequately, Not at All	25.3%	18.6%	24.8%	34.8%
No Opinion	5.1%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1364

Q14(c) To what extent is it appropriate to award exceptional quality points for each indicator and for the whole criterion?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	38.2%	21.4%	24.1%	63.8%
Inadequately, Not at All	55.5%	78.6%	75.9%	36.2%
No Opinion	6.3%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1347

These questions generated some of the more significant differences between administrators and teachers. The current system requires that administrators document the observance of EQs; teachers must earn EQs in order to obtain higher scores. Given this background, it may not be surprising to see such differences of opinion. In question Q14(b), there was moderate to strong support at all positions for applying the EQ at the criterion level. However, when asked whether EQs should be applied at the indicator level or at the indicator level in conjunction with the criterion level, teachers were neutral or moderately supportive. Administrators were strongly against this practice, which would, of course, require extra documentation. During the first year of the TTAS, EQs were awarded at the indicator level, but based on recommendations from an external committee, this practice was changed.

On-Site:

Results from the on-site interviews indicated general agreement with the responses from the questionnaire. Additional comments included suggestions from some respondents that EQs be abolished. Some of the teachers stated that EQs should be based on overall and cumulative quality of teaching, while others suggested awarding EQs only to those who exceed a clearly defined baseline of behavior.

Open-Ended Comments:

(Principal) Do away with documentation for EQ. Spend time documenting areas targeted for growth. This would be more productive and improve teaching.

(Teacher) Take the mystery out of EQs by clearly defining EQ behaviors in TTAS. Make appraisers more responsible for raising teacher scores with feedback appropriate to individual classrooms.

PEIMS Data:

Although there are no data in PEIMS that directly address the issue of the EQs, unexpected breaks in the distributions of the scores, to be discussed in detail later, might be related to EQs, since there was a six point difference between some of the drop points.

Communication and Observation

Teacher/Appraiser Interactions. The next four questions specifically concern the operation of the system in the direct interaction between appraiser and teacher. Pre- and post-conferences are vehicles for communicating exact expectations of behavior and discussing observations which lead to the assessment score. Appropriate feedback is critical to the success of any appraisal system, regardless of profession. Also examined is the length and frequency of observations. Clearly, there must be a tradeoff between assurance of accurate and thorough observation and intrusion into the instructional process. This is especially true for scheduled visits where teachers are much more likely to prepare lessons for the benefit of the appraiser rather than for the benefit of instruction on that particular day.

Survey Questions:

Q15(a) To what extent is the use of pre-conferences appropriate?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	71.2%	74.9%	76.0%	68.0%
Inadequately, Not at All	27.1%	25.1%	24.0%	32.0%
No Opinion	1.6%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.010 df=2 n=1416

Q15(b) To what extent is the use of post-conferences appropriate?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	86.7%	92.6%	92.7%	81.2%
Inadequately, Not at All	12.4%	7.4%	7.3%	18.8%
No Opinion	0.9%	n/a	n/a	n/a

By Position: Statistically *Significant*; Chi-square F<.001 df=2 n=1425

Q15(c) To what extent is the length of time periods of observations appropriate?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	74.3%	76.9%	73.6%	73.8%
Inadequately, Not at All	24.4%	23.1%	26.4%	26.2%
No Opinion	1.2%	n/a	n/a	n/a

By Position: Statistically Non-Significant; Chi-square F<.411 df=2 n=1422

Q15(d) To what extent is the frequency of observation appropriate?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	62.9%	64.8%	59.9%	64.3%
Inadequately, Not at All	35.4%	35.2%	40.1%	35.7%
No Opinion	1.8%	n/a	n/a	n/a

By Position: Statistically Non-Significant; Chi-square F<.361 df=2 n=1414

There was obvious universal support for both pre- and post-conferences, with the post-conferences receiving exceptional support. Although the comments noted on the survey have a different flavor, the length of the time period for observations received moderate support from all respondents. There was moderate, but somewhat less enthusiastic, support for the frequency of observations. It should be noted that even though it is allowed under current rules, the use of informal observations which might lead to a modification of awarded scores was not identified as an issue prior to the survey. In addition, no comments were received concerning this practice, indicating either little concern regarding this option or infrequent use of it.

On-Site:

The comments from the on-site interview indicate that the pre- and post-conferences were viewed as helpful by all respondents. The length of observation time period was also generally supported. Teachers were much less positive about the frequency and length of observation periods than were the superintendents and principals.

Open-Ended Comments:

(Superintendent) Four announced visits makes the system ineffective and unappreciated. A sound system of routine classroom visits by the principal in addition to TTAS would result in a much better plan.

(Principal) Scheduled observations only provide an opportunity for teachers to put on a one hour show 2 or 4 times a year. It doesn't recognize the teacher who does an outstanding job for each class every day.

(Principal) I believe that several informal observations would be more beneficial than 2-4 formal observations.

(Principal) I feel that scheduled visits have changed the ability for supervisors to accurately observe teaching. There are also observable indicators that need to be added to the instrument. Often it is difficult to "fit" certain areas of classroom management into the instrument in its present form.

(Teacher) The TTAS only measures a specific teaching time, 45 minutes, from 1 to 4 times a year. I have seen extremely poor teachers rated outstanding because they want to perform only during that time. I have seen great teachers "blow it" and be penalized because they were nervous or because of a disruptive student. This instrument only tests a 45 minute segment, be it good or bad and does not weed out poor, ineffective teachers.

(Teacher) Change the frequency so that proven, quality teachers are evaluated every 3 years.

Accuracy and Problems

Accuracy of Observation. Another question addressed whether the TTAS allows accuracy of observation.

Survey Questions:

Q15(k) Are appraisers capable of using the TTAS accurately as it is presently designed?

	Overall	Principals	Superin-tendents	Teachers
Completely or Adequately	72.2%	81.7%	78.0%	66.0%
Inadequately, Not at All	25.4%	18.3%	22.0%	34.0%
No Opinion	2.4%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1406

There was very strong support from the administrators and teachers for the capability of the TTAS to be used accurately. As noted previously, capability and actual use may, of course, differ.

Locus of Difficulty. The next question combined the concept of the TTAS itself and the mechanics of implementation. There were four choices presented that were designed to focus the question of difficulties on either the TTAS or its implementation in a mutually exclusive fashion.

Survey Questions:

Q18 Which of the following statements most accurately reflects your opinion of the relationship between the TTAS and its implementation?

1. The major difficulty is with the TTAS, not how it is implemented in my district.
2. The major difficulty is with the way the TTAS is implemented in my district, not with the TTAS itself.
3. The major difficulty is with both the TTAS and the way it is implemented in my district.
4. There is no major difficulty with either the TTAS or the implementation in the district.

	Overall	Principals	Superin-tendents	Teachers
TTAS, not implementation	33.2%	37.3%	32.2%	30.6%
Implementation, not TTAS	17.1%	16.1%	22.4%	15.9%
Both	22.6%	11.6%	16.3%	33.8%
Neither	27.1%	35.0%	29.1%	19.7%

By Position: Statistically Significant; Chi-square F<.010 df-2 n-1384

Approximately one-third of administrators indicated there are no problems with either the TTAS or the manner in which it is being

BEST COPY AVAILABLE

implemented. Conversely, about one-third of the teachers responded that there is a problem with both. This reflects the difference in orientation described earlier between evaluators and those being evaluated. In any case, a majority of both administrators and teachers believe that there are problems with either the TTAS or its implementation.

On-Site and Open-Ended Comments. In a non-guided question asking what the major difficulty is with the TTAS in the district, the most frequent response was bias or subjectivity by the appraiser. This topic will be examined in depth in the third chapter of the results. No significant new information was obtained from written comments on the survey.

Discussion and Recommendations:

There is a very clear difference of opinion between administrators and teachers concerning exceptional quality points. While both groups favor continuing the practice of awarding EQs at the criterion level, teachers tend to support the idea of awarding EQs at the indicator level as well. Administrators were very opposed to this idea, possibly due to the potential for increased documentation requirements for each award. One possibility to address this issue is to explore awarding EQs at the criterion level and reducing the amount of documentation required. It remains to be demonstrated that this would increase the dispersion of scores. Simply raising the total possible score without a concomitant increase in variance among teachers' appraisal scores will not increase the usefulness of the TTAS.

There is a great deal of support for conferences between teacher and appraiser. Because an improvement plan must be written only when scores are less than satisfactory and almost none are, the conference may be the only means of formally discussing suggestions and comments. Again, since the lower categories are almost never used and the middle category is rarely used, formalized expectations for improvement may not be a viable part of this system as it currently exists.

Although there was support for the capability of appraisers to accurately use the TTAS, there was indication that major problems exist. The majority of respondents believe that there are problems with either the TTAS or its implementation. By combining categories of responses, it appears that 64.4 percent of teachers and 48.7 percent of administrators believe that there are major problems with the TTAS. Additionally, 49.7 percent of teachers and 34.6 of administrators believe that there are problems with implementation. It might be tempting to say that there are fewer problems with implementation than with the system itself, but the percentages of educators who believe there are problems are fairly high in either case.

The accuracy of appraisals can be more readily addressed than can fundamental problems with the appraisal system itself. According to Edwards (1989) and Fox (1987/88) the use of multiple sources and multiple raters increases the accuracy of appraisal. Lupton (1989) indicates that reducing the number of categories of performance to those that supervisors can reliably distinguish and clearly explain also improves accuracy.

Recommendation 2.1. Establish a pilot program in selected districts under careful controls to examine awarding EQs at the indicator level.

An experimental program would allow examination of the effect of awarding EQs at the indicator level. The amount of time, explanation, and documentation required can be evaluated. It is assumed that the addition of more EQs will increase the range of possible scores on the TTAS. If the variance of scores among teachers does not increase however, then the discrimination power of the TTAS will not increase, making the addition of EQs at the indicator level of questionable value. This recommendation might be combined with recommendation 2.2.

Recommendation 2.2. Establish a pilot program in selected districts under careful controls to examine documentation of EQs. Restriction, enhancement, or removal of requirements for documentation should be examined.

This recommendation would look at the impact, if any, of changing the amount or type of documentation required for awarding EQs.

Recommendation 2.3. After the first two years of teaching, eliminate annual evaluations. Appraise every third year with a minimum of four observation periods, only one of which would be scheduled.

This recommendation draws on information obtained from questions used in this study, teacher and administrator comments, TTAS score distributions, and teacher performance classifications. Although there is agreement that the TTAS is suited for basic teaching skills, there is doubt about its effectiveness at higher level instructional activities. There is general concern that the scheduled visits are "dog and pony" shows which may have little correlation to normal instruction. There is no doubt that preparation for these visits is time-consuming and anxiety-producing. Keeping the observation periods at the same length, but increasing the number of visits during the year will help insure adequate observation and, perhaps, enable the observers to better understand the particular teacher. Requiring observations during the first two years, possibly with more frequency, would help ensure acceptable levels of performance for beginning teachers.

It should be noted that this recommendation is predicated on an increase in the validity of assessment from the TTAS. If there is doubt as to initial classifications of teachers, then removing teachers from formal assessment of their performance for an extended period of time is inappropriate. This recommendation must be considered in light of findings from Recommendation 1.1 concerning student performance.

Results

Chapter 3: Subjectivity or Bias

A standardized statewide evaluation procedure using a structured instrument is certainly expected to reduce the potential for personal bias or subjectivity to influence ratings received by teachers. In addition to subjectivity, a potential source of bias can be lack of specific knowledge by the appraiser, such as when the appraiser is not familiar with the instructional techniques or subject area being observed. Given the importance of this topic, several questions within the survey were written to directly assess it. In addition, other questions within the survey reflected the theme of bias or subjectivity. In fact, many questions within the survey elicited responses which are applicable across several of the five major issues. Therefore, responses to some questions may be included in more than one chapter.

Survey Questions:

Q8(a) In your district, to what extent does the TTAS reflect actual teaching ability or competence?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	57.7%	72.2%	59.8%	46.6%
Inadequately, Not at All	41.5%	27.8%	40.2%	53.4%
No Opinion	.8%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1428

Q8(g) In your district, does the TTAS objectively and impartially appraise teacher performance?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	45.7%	60.3%	50.2%	33.5%
Inadequately, Not at All	53.3%	39.7%	49.8%	66.5%
No Opinion	.9%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1425

Q8(h) In your district, does the TTAS result in more fairness of teacher appraisal than earlier evaluation practices?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	54.2%	71.8%	66.7%	43.3%
Inadequately, Not at All	38.0%	28.2%	33.3%	56.7%
No Opinion	7.8%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1326

There were considerable differences in opinion regarding these statements among the various professional positions. Although each of the comparisons was statistically significant, the reader is again cautioned that the large number of respondents to these items makes statistical significance almost a foregone conclusion. As shown in the tables, principals expressed strong support for the TTAS in terms of its capability to reflect actual teaching ability and provide fairness of appraisal, and expressed moderate support for its objectivity. Superintendents, while somewhat more positive than teachers, were more likely to be neutral or moderate in their support for these three positions. Teachers were neutral in their positive support for fairness and reflection of teaching ability by the TTAS and indicated moderate non-support for the statement that the TTAS system is objective and impartial. It is interesting to note that, overall, Q8(h) concerning fairness had almost eight percent of respondents who chose not to respond. Considering the importance of this issue, it is somewhat surprising that such a relatively large percent chose not to respond or did not have an opinion concerning the fairness of the TTAS.

From these three questions, it can be stated that there was support, or at least neutral opinion, for the ability of the TTAS to reflect actual teaching ability and to be fairer than previous practices. There was much less support, and in fact, some opposition on the part of teachers that the TTAS objectively and impartially appraises teacher performance. In order to examine this issue further, three additional questions were considered which address specific features of the TTAS designed to decrease subjectivity by defining certain teacher behaviors. These features are especially applicable to the situation where the appraiser may not be familiar with the techniques being employed by the teacher.

Q12(c) To what extent does the TTAS clearly describe or define observable teaching behaviors?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	77.3%	86.8%	84.7%	68.4%
Inadequately, Not at All	21.7%	13.2%	15.3%	31.6%
No Opinion	1.0%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1424

Q12(d) To what extent does the TTAS clearly describe or define quantifiable teaching behaviors?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	64.3%	76.1%	65.9%	57.1%
Inadequately, Not at All	33.4%	23.9%	34.1%	42.9%
No Opinion	2.3%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1406

Q12(e) To what extent does the TTAS clearly describe or define the quality of teaching performance?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	54.2%	58.4%	50.9%	47.3%
Inadequately, Not at All	38.0%	41.6%	49.1%	52.7%
No Opinion	7.8%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1422

In contrast to the concept of fairness explored in the previous set of questions, there appeared to be strong support that the instrument describes or defines observable teaching behaviors. As described in Chapter 1, the question concerning quantifiable teaching behaviors evoked strong support from principals, moderate support from superintendents and more neutral support from teachers.

The picture is not so clear when asked about the quality of teacher performance. In this case all three groups indicated support in the neutral range. It is clear that teaching is a complex task and should not be construed as a simple summation of observable behaviors which can be labeled effective instruction. Teaching, rather, may be viewed as the integration and use of many abilities and actions which should lead to enhanced student learning. The respondents' opinions were clearly split on the ability of the TTAS to address this issue and it is interesting to note that the response patterns for all position types were similar. As with the first set of questions, almost eight percent of respondents either did not have an opinion or chose not to indicate one for these questions.

Since there was disagreement among the professional positions concerning the objectivity, impartiality, and fairness of the TTAS, yet general agreement on the ability of the TTAS to describe observable and quantifiable behaviors, this led to questions regarding another aspect of the TTAS system. Thus, results from questions regarding training of appraisers were explored.

Q15(e) To what extent is the monitoring of appraiser procedures appropriate for the appraisal of effective teaching practices?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	63.7%	78.8%	64.9%	64.5%
Inadequately, Not at All	27.0%	21.2%	35.1%	35.5%
No Opinion	9.3%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1306

Q15(f) To what extent is the quality of training and appraisers appropriate?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	74.3%	85.2%	83.2%	72.3%
Inadequately, Not at All	19.2%	14.8%	16.8%	27.7%
No Opinion	6.5%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df-2 n-1344

Q15(g) To what extent is the frequency of training for appraisers appropriate?

	Overall	Principals	Superin-tendents	Teachers
Completely or Adequately	73.7%	88.3%	82.9%%	76.4%
Inadequately, Not at All	15.6%	11.7%	17.1%	23.6%
No Opinion	10.7%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1283

From the results presented above, it is clear that there was considerable support across all types of respondents for the training and monitoring of appraisers. Continuing to search for the locus of uncertainty leading to the concern over fairness, three questions concerning the abilities of appraisers were examined.

Q15(h) How consistent are the interpretations of the TTAS indicators by appraisers?

	Overall	Principals	Superin-tendents	Teachers
Completely or Adequately	56.5%	63.4%	65.1%	49.5%
Inadequately, Not at All	41.1%	36.6%	34.9%	50.5%
No Opinion	2.4%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1403

Q15(i) Is the specific subject area knowledge of appraisers appropriate as used within the TTAS?

	Overall	Principals	Superin-tendents	Teachers
Completely or Adequately	57.1%	67.1%	57.2%	51.8%
Inadequately, Not at All	40.4%	32.9%	42.8%	48.2%
No Opinion	2.4%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1403

Q15(j) Is the pedagogical knowledge and experience of the appraisers appropriate as used within the TTAS?

	Overall	Principals	Superin-tendents	Teachers
Completely or Adequately	69.8%	78.1%	77.8%	66.9%
Inadequately, Not at All	25.5%	21.9%	22.2%	33.1%
No Opinion	4.7%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1344

Although there was strong and moderate support for the premise that appraisers' pedagogical knowledge and experience is acceptable, there was less endorsement for the specific subject area knowledge of the appraiser with only moderate support from the principals (who are often appraisers), and a neutral stance from the teachers and superintendents. According to Berliner (1987, pp.39-65), the pedagogical knowledge level of the appraiser is critical for consistent ratings. Several studies (Hahn & Dipboye, 1988; Heneman, Wexely, & Moore, 1987; Kamouri & Balzer, 1990) indicate the need for evaluators to have specific information about the job to be rated.

With regard to consistency of interpretation of the TTAS indicators, principals and superintendents evidenced moderate support while teachers were neutral, with half of the respondents on either side of this issue. Heneman, et al., (1987) indicated that consistent appraisals were associated with appraisers who believe in the variability of human nature, have a behavioral frame of reference, and have enhanced opportunities to observe performance. In a comparison of appraisers with differing levels of training, Berliner (1987) noted that even "advanced beginners" (using a definition specific to his study which is not necessarily applicable to the current TTAS appraisal process) reported contradictory interpretations of the same teaching and learning events. Clearly, literature indicates that extensive training, continued practice, and objective definitions of behaviors are needed to ensure consistent interpretations. The issue of a need for a cadre of "professional" appraisers remains unanswered in this study.

An additional set of questions [Q13(a) through Q13(f)] directly addressed the issues of subjectivity, bias, and other potential influences. However, difficulties with interpretation of these questions because of the need to reverse the scale (i.e., "completely" implied negative impact) when completing the questionnaire were noted by respondents. A conference including program staff from the agency, the outside contractor, and evaluation staff concluded that the responses to these questions could easily have been misinterpreted. For this reason, results from these items will be not be presented.

On-Site Interviews:

In general, the interview results supported the data obtained from the questionnaire. For example, only 19 percent of the interviewees felt that the TTAS was subjective, 14 percent stated that appraiser bias affects the scores, and 24 percent of the teachers noted the presence of appraiser bias.

To the direct question "Do you believe that TTAS scores are influenced by factors other than teacher performance?" an overwhelming majority of respondents (89 percent) said "YES." Forty-nine percent then stated that the TTAS was subjective, a result which is quite interesting since earlier in the interview, only 19 percent made this same observation. Clearly the direct question investigating influences on scoring evoked stronger feelings than earlier, indirect questions. This question followed three inquiries regarding 1) observable teaching behaviors, 2) quantifiable teaching behaviors, and 3) quality of teaching performance. These three questions evoked responses similar in pattern to those found from analysis of the questionnaire.

Open-Ended Comments:

(Teacher) Outside appraisers would make the current system more objective.

(Teacher) My main concern is the great inconsistency of interpretations by appraisers. It is very helpful to be appraised by on-campus evaluators who have some knowledge of the area of behavior of students in the classroom.

(Teacher) The instrument is more subjective than ever. Appraisers are allowed too much bias in interpreting quality teacher performance. It does not address student accountability for learning.

(Superintendent) Major problems are related to lack of knowledge by administrators about effective teaching practices, whether or not learning is occurring, being able to recognize need for improvement, etc.

There were also many comments linking subjectivity or bias to use of the TTAS to support the career ladder. This issue will be addressed later in this report.

Data from PEIMS:

The PEIMS database contains a variety of information collected from all public schools in the state. For this part of the evaluation, appraisal scores were examined in relation to a variety of teacher characteristics to look for the presence or absence of bias, or targeting of scores.

Patterning. The first step in analyzing these data was to look for recurring patterns in the scores. Given a very large number of scores, the expectation is that the distribution of scores will assume some variation of a normal curve. Although not all curves of this nature are symmetrically distributed, they usually will retain many important aspects of normal curves, one such aspect being a relatively smooth shape. If a distribution of a large number of scores is examined and significant deviations from the general shape are seen, then some underlying factor can reasonably be suspected. Further, if numerous measurements taken at repeated intervals produce repeated occurrences of the deviant pattern, it is probable that some external influence is present.

Data were available from four measurement periods: fall and spring - school year 1988-89, and fall and spring - school year 1989-90 for approximately 180,000 Texas teachers. Patterning was apparent in all these distributions. An example of this patterning, obtained from the frequency distribution of the appraisal scores reported in the spring of 1990, is presented in Table 3. These scores are from the upper range of the distribution.

TABLE 3
Deviation From Normal Curve

Total Score	Number of Teachers With This Score
. . .	
170	4637
171	3912
172	4017
173	3135
174	3273
175	642
176	2800
177	2591
178	2239
179	1874
180	2022
. . .	

As the distribution of appraisal scores approaches the maximum possible score of 184, the normal curve expectation is that the number of teachers with a particular score would start to diminish. In fact, this pattern does appear. Notice, however, the substantial drop in the number of teachers with a total score of 175. As stated previously, such a deviation would not be remarkable given one measurement with the attendant possibility of random effects. However, similar large drops at this specific score were noted in each of the four measurement intervals over a two year period. In addition to the remarkable drop in frequency at this score point, there were isolated occurrences of "dips" at other points in each of the four score distributions. The rule used to classify these deviations was that they had to occur at least three times out of the four measures over two years. The table below lists those total scores at which a deviation from a smooth pattern was marked by at least a 50 percent drop in frequency from the previous score. The majority of those deviations included in the table exceeded this difference from the distribution curve. Scores at the lower end of the distribution were not included due to the relatively smaller numbers of teachers at this score range and the resultant potential instability of the data. Scores excluded from the table were those below a total score of 100 and represent less than 1.6 percent of all scores. The deviation from the pattern at score 104 was a significant increase, rather than a decrease. Table 4 also presents the categorical ratings associated with score ranges.

TABLE 4
Points in Total Score Distribution With Deviations From Pattern

Total Score	Number of Occurrences (out of four measurements)
92	3
104	3
139	3
145	4
151	4
157	4
163	4
169	4
175	4
181	4
183	4

REFERENCE TABLE

TTAS Rating Category	Total Score Range
Unsatisfactory	below 79.9
Below Expectations	80-103.9
Meets Expectations	104-135.9
Exceeds Expectations	136-159.9
Clearly Outstanding	160-184

In addition to these deviations, there was a significant bulge in the number of teachers at the highest score of 184 (3756 teachers at 184, 379 at 183, and 1922 at 182). In fact, this number represents such a large deviation that in order to find more teachers at a lower score, you must drop down in the distribution to a total score of 172. There are, of course, more teachers with particular scores lower down the scale, reflecting the normal distribution of total scoring.

While it would have been helpful from an explanatory point of view for the deviances to occur exactly at the score breaks between rating categories, this was not the case except at score 104. It was at score 104, the break between Below Expectations and Meets Expectations, that there was an increase in the frequency of scores. This pattern may not be remarkable given a targeting of scores to indicate teachers just barely meeting expectations. Otherwise, it is not clear what is driving these deviations, only that they are occurring. It is clear, however, that the appearance of non-random fluctuations in score distributions is indicative of external influences upon the scores. It is recognized that the publication of this information may, in fact, cause a new external influence upon the score distributions in the future.

Percent of Teachers Classified in Various Categories. Another method of examining external influences is to look at the overall distribution of classifications. Certainly, few people relish the label of "average" or "satisfactory", regardless of profession. The TTAS attempts to avoid a label of average, but when the label "meets expectations" is surrounded by "exceeds expectations" and "below expectations," clearly the concept of average is inherent to the middle category. It should be noted that the label for this middle category was changed from "satisfactory" to "meets expectations" following the 1988-89 school year. Although this report does not examine the philosophy of selecting particular labels for "middle" or "expected" performance, if expectations are appropriately set for average performance then, by definition, reasonable measurement will place most outcomes and people in this range. By most definitions most people are average in their performance, even though this is clearly not a classification most people willingly accept.

The number and percent of teachers who fall into the various TTAS categories are given in Table 5. Since examination of data over the four measurement periods reveals similar patterns, the results from the most recently available data collection (spring, 1990) are presented below:

TABLE 5
Number and Percent of Teachers in Various Categories

Category	Number	Percent	Score Range
Unsatisfactory	37	0.0%	below 79.9
Below Expectations	257	0.2%	80-103.9
Meets Expectations	11,771	6.6%	104-135.9
Exceeds Expectations	82,447	46.2%	136-159.9
Clearly Outstanding	83,761	47.0%	160-184
TOTALS	178,273	100.0%	

Regardless of efforts to build construct validity into the TTAS, it is apparent that the vast majority of teachers are classified as "exceeds expectations" and "clearly outstanding." The overall academic standing of many Texas schools causes the validity of these ratings to be suspect.

Average (Mean) Scores. Another way of analyzing the classification of teachers is to examine the mean appraisal scores over the four measurement periods. Mean scores aggregated at the teacher level for the four time periods used in this analysis are presented in Table 6.

TABLE 6
TTAS Mean Score For All Teachers

Fall 1988	Spring 1989	Fall 1989	Spring 1990
153.4	153.0	154.7	157.4

The scores, while showing an increasing trend, are all in the "exceeds expectations" range; for the spring of 1990 the mean score was very close to the "clearly outstanding" category. Again, this very high mean, as a reflection of teacher performance, is not consistent with current student performance levels.

Repetitive Scores. Yet another method of examining deviations from expected measurement of human behavior is to analyze score repetitions.

BEST COPY AVAILABLE

As was clear in Table 4, the majority of teachers are classified in one of two categories. Even considering that an ostensible two-category system will be indicative of a restricted range of scores, the odds of a single teacher receiving exactly the same score over four different appraisal periods during two different years with, in many cases, different appraisers are extremely small. Although some classroom observations are scheduled, which can lead to "staged" instructional sequences, teacher behavior does vary, and observers will have some inconsistencies in perception.

A not yet published study conducted by Barnes (1990) under very controlled circumstances indicates a substantial variation in scores assigned to teacher behaviors by a small number of carefully trained appraisers. The range of differences in average scores given teachers by appraisers varied from 3.33 to 14.66 for one appraiser team and from 2.34 to 12.67 for another team. The study states "Again it is clear that teacher within campus scores varied widely across appraisers." Further, the study concludes "Teacher within campus scores were significantly affected by factors other than teacher skill." What is apparent is that even given a neutral setting with extensively trained appraisers, a substantial amount of variation of teacher scores is to be expected.

To examine the possible occurrence of repeated scores, a sample of 5,000 teachers' appraisal scores was drawn from the PEIMS database. It was found that seven percent of the teachers in the sample had exactly the same score over the four measurements available. Additionally, fully 21 percent of teachers had the same totals within two score points. While the assumptions concerning restriction of range make exact calculations of the probability of these occurrences problematic, the expected fluctuations in teacher and student behaviors and observer perceptions should produce odds ranging from astronomical to unbelievable - without outside, non-random influence. According to Phillips (1987), "red flags" indicating a problem with a performance appraisal system include:

- identical performance ratings for all employees and
- identical performance ratings and comments for the same employee from one appraisal period to the next.

Demographics of Districts. An analysis tool available within the agency allows scores to be examined by a number of demographic measures at the district level. Appraisal scores across the four measurement periods reveal few apparent relationships with district demographics. The tables generated are included in Appendix E; the following statements result from examination of the patterns in these tables.

No consistent, strong relationship exists among TTAS scores and:

Wealth per Student
District Size
Tax Effort
ESC Region
Percent Minority Students

Percent Low Income
Percent Minority Teachers
Teachers with Advanced Degrees
Average Teacher Experience.

Some relationship may be evident among TTAS scores and:

Percent Passing All Tests Taken on TEAMS
Average Teacher Salary.

These findings are interesting in light of the possibility that TTAS scores may be restricted in some districts in order to reduce the number of teachers eligible for the career ladder. While mean TTAS scores did increase somewhat with teacher salary, they did not increase with increasing wealth per student. The range of average scores by teacher salary was quite small, from 152.6 for the lowest average teacher salary group to 158.7 for the highest average salary group of districts. Both of the averages were in the upper range of the "exceeds expectations" category of the TTAS. Overall, it is not apparent that any of the district demographic characteristics are a factor in determining average TTAS scores. The percent of students passing all tests taken on the Texas Educational Assessment of Minimum Skills (TEAMS) will be examined in a later section on the impact of the TTAS. This issue will be further discussed in the last chapter of the results section when the career ladder is examined.

Other Bias. Another form of bias would be evident if specific groups of teachers received lower scores. In order to examine this possibility, Table 7 presents summary information available from the PEIMS data base.

TABLE 7
Total TTAS Score by Teacher Demographics

	Percentage of Teachers by Demographic Within Score Range					
TTAS Score Range	Sex		Ethnicity*			Label Range
	Male	Female	Black	Hispanic	White	
32 < 136	10.1	5.0	8.5	7.7	5.5	Unsatisfactory, Below/Meets Expectation
136 < 146	20.3	12.7	15.7	16.2	13.8	Exceeds Expectation
146 < 154	20.7	16.9	16.8	18.6	17.6	
154 < 160	15.5	15.2	13.6	15.8	15.4	
160 < 170	22.8	29.2	27.1	27.2	28.0	Clearly Outstanding
170 <=184	10.7	21.2	18.4	14.5	19.8	
Totals	100.0	100.0	100.0	100.0	100.0	

* American Indians and Asians are omitted due to very small numbers.

From an examination of Table 7, it appears there is little evidence of a pattern of bias when total score is examined across ethnic groups, since the percentages are very close to each other. However, the pattern is different for male versus female. Approximately half of the females receive scores above 160 (clearly outstanding) while only one-third of the males are in this category. A similar analysis is presented in the last chapter for the distribution of career ladder assignments.

There are, of course, some patterns that would be expected when years of experience and level of education are considered. Each of these are included in the criteria for advancement on the career ladder and each could reasonably be expected to have some relationship to outstanding teacher behaviors. Table 8 presents these results.

TABLE 8
Total TTAS Score by Experience and Degree

Percentage of Teachers by Demographic Within Score Range

TTAS Score Range	Years of Experience			Degree			Label Range
	0-5	6-10	11-15*	Bachelor	Master	Ph.D	
32 < 136	10.2	5.2	4.1	6.9	4.3	5.8	Unsatisfactory, Below/Meets Expectations
136 < 146	20.4	14.0	11.6	15.8	11.2	12.0	Exceeds Expectations
146 < 154	22.2	18.3	15.8	19.2	14.7	15.0	
154 < 160	16.2	15.8	15.3	15.8	14.3	11.8	
160 < 170	21.0	28.3	30.6	25.9	31.5	29.6	Clearly Outstanding
170 <=184	10.0	18.4	22.5	16.4	24.0	25.9	
Totals	100.0	100.0	100.0	100.0	100.0	100.0	

*The patterns do not change from 11-15 through 30 years plus of experience.

There is an expected increase in TTAS score after the first grouping of years of experience, which might reflect an increase in teacher skills after obtaining classroom experience. Another hypothesis is that teachers are better able to prepare for an appraisal observation after a certain amount of experience. There is a relationship between earned degree and score, with the advantage most evident in the clearly outstanding range. Unfortunately, it is impossible to determine from these data a causal relationship. That is, more experienced, higher-degreed teachers may be better instructors who earn higher TTAS scores, or the *a priori* expectations of the appraisers, given knowledge of the teacher's credentials, may be higher, leading to enhanced scoring.

Discussion and Recommendations

According to published literature, essential requirements for implementing an effective performance appraisal system include the characteristics that performance appraisal is objective (Webster, 1988), is based upon a clear description of expected performance in terms of specific behaviors rather than vague attributes (Pelle & Greenhalgh, 1987), and is not used to justify preordained decisions (Webster, 1988).

While the opinions of administrators were that the TTAS does indeed reflect actual teaching ability and is fairer than previous systems, teachers did not share this perception. When questioned about the objectivity and impartiality of the TTAS, teachers clearly question the system and administrators were less sure than with the previous two questions. Respondents agreed that observable and quantifiable teaching behaviors are well described and defined. There was less agreement that the quality of teaching is measurable with this system. From this information, it is evident that educators feel the TTAS reflects isolated teaching skills and contains needed definitions, but that scores may not be objectively determined. Brandt (1990) asserts that problems of educational rating practices concerning equity and fairness include evaluation bias, halo tendencies, and inflation of scores.

There was little indication of a problem with monitoring or training of appraisers and there was agreement that interpretations are fairly consistent, reflecting adequate training. The pedagogical knowledge of appraisers was not questioned, but there was more likely to be a neutral, or split, opinion regarding the subject area knowledge for appraisers. Pedagogical knowledge acquired in a university setting is more likely to be reinforced in appraisal training sessions than is specific subject matter.

It is evident from examination of PEIMS data that patterning and repetition of scores were consistent with a premise that influences outside the TTAS are driving the scores. From earlier discussions it was noted that substantial dips in score frequencies occurred with regularity. Even though an explanation is not proposed for this pattern, it is clear that the strength and regularity of the deviations in score patterns make some external influence certain.

Overall, mean appraisal scores for teachers were quite high and only two classifications were actually used - "exceeds expectations" and "clearly outstanding." Problems should be evident in a system where there is an inflation of scores (Brandt, 1990). Further, a very large percentage of teachers received identical scores over several observation periods and, it is assumed, with different appraisers. Although the high percentage of teachers with essentially the same scores over multiple time intervals might argue for reliability of the system, in fact, over-reliability is a clear indication of appraiser bias (Phillips, 1987). The TTAS, as practiced, appears to be so over-reliable that its validity is highly questionable. A system where there is no effective discrimination among teachers' abilities has virtually no utility.

No matter how carefully an observation system is designed in terms of standardization and objectivity, there always exists the potential for appraiser bias. Two external forces which might influence scores on the TTAS are immediately apparent. The first is the inherent correspondence between TTAS scores and career ladder decisions. Given the need for high levels of performance on the TTAS in order for teachers to move up the career ladder, there exists a pressure for upward bias in the scores. Appraisers are likely to be reluctant to produce unfavorable

TTAS ratings which would then preclude career ladder bonuses to teachers who, it is generally agreed, are underpaid.

The second force driving TTAS scores upward is the reluctance to assign a person to the category "meets expectations" since this is equivalent to average performance. Given the educational context of the TTAS, there is also an implicit link between the five rating categories on the TTAS and the traditional five grades given to students: A, B, C, D and F. The inability to accept that one's performance rates a C, or average score, is a universal coping mechanism which ensures a sturdy level of self-esteem. However, most human behavior is, by definition, average. Until, or unless, it can be accepted that "average" or "meets expectations" is an acceptable state of affairs, there will be pressure for higher scores.

According to Geber (1988), it is not possible to design a performance appraisal system that ensures perfect objectivity. Further, she asserts that no matter how much a company may spend creating a system and providing training, managers and supervisors still tend to falsify performance ratings by deflating or inflating an employee's ratings. Managers and supervisors tended to place more importance on the use of appraisals as a way to motivate employees rather than on rating accuracy. Reasons for deflating appraisals include the desire to encourage an employee to quit, and to create a record to justify a planned action such as a demotion or firing. Reasons for inflating appraisals include the desire to help a worker whose performance is suffering because of temporary personal problems, to avoid creating a permanent record of poor performance and to avoid confrontation with hard-to-manage employees.

Recommendation 3.1. Disassociate the TTAS from the career ladder.

Implementation of this recommendation would diminish the pressure for exceptionally high scores by dissociating the TTAS from decisions regarding salary supplements. The removal of the association between the TTAS and the career ladder will also help to remove any artificial caps on scores which might be driven by limited career ladder funds.

Recommendation 3.2. Evaluate appraisers based on a distribution of scores awarded over time, especially as compared to statewide information. Publish means and distributions of campus, district, and statewide TTAS scores.

Publication of statewide information will help determine the extent of problems with scoring and identify trends toward more reasonable scores. Requiring appraisers to examine their distribution of scores and provide justifications for scores awarded as part of their evaluation process should make scoring more reasonable.

Recommendation 3.3. Continue the educational process for appraisers, teachers, and the general public concerning the meaning and utility of the various rating categories. Universities associated with teacher training need to strengthen curriculum in this area.

Learning how scores should be used and what meaningful expectations exist for distributions of scores will help diminish the feeling that average or satisfactory performance is an anathema. Perhaps an investigation into the use of other terms and numbers of levels would be helpful as well.

Recommendation 3.4. Investigate the use of a cadre of "professional" appraisers, perhaps based at the education service centers.

This particular recommendation would examine the impact of using personnel whose only job is to conduct appraisals. This should, of course, ensure familiarity with pedagogical knowledge and the TTAS, but might reduce familiarity with specific subject matter knowledge. The use of this type of appraiser either exclusively or in conjunction with "regular" appraisers should be investigated.

Results

Chapter 4: Uses of the TTAS

In addition to assessing teacher effectiveness, potential applications of the TTAS include providing information to be used in conjunction with other requirements for placement on the career ladder, for contract renewal decisions, and for campus assignments. Each of these uses will be explored in this chapter. One question specific to the future of the career ladder is examined in the final chapter of the results section.

Career Ladder and Master Teacher Designations. As currently practiced in Texas, the career ladder cannot be dissociated from the TTAS because of the legislative mandate detailed in Appendix A of this report. Thus the results of the following question were of interest:

Survey Questions:

Q16(b) To what extent should the TTAS be linked to career ladder decisions?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	49.0%	55.7%	42.9%	48.0%
Inadequately, Not at All	49.3%	44.3%	57.1%	52.0%
No Opinion	1.7%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.003 df=2 n=1415

Somewhat different from the typical response pattern found in the analyses of this questionnaire, the teachers were more supportive of this idea than were the superintendents. Given the fact that opinions were clearly split with almost equal percentages on both sides of the issue, the complete disassociation of the TTAS and the career ladder cannot be supported from the results of this questionnaire. The next four questions yielded a greater divergence of opinion.

Q11(a) To what extent is the use of the present TTAS instrument appropriate for appraisal of master teachers?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	41.9%	53.7%	35.9%	42.1%
Inadequately, Not at All	50.7%	46.3%	64.1%	57.9%
No Opinion	7.4%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1334

BEST COPY AVAILABLE

Q11(b) To what extent is the use of a different scoring procedure with the same TTAS instrument appropriate for appraisal of master teachers?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	24.4%	32.1%	27.8%	26.5%
Inadequately, Not at All	61.2%	67.9%	72.2%	73.5%
No Opinion	14.4%	n/a	n/a	n/a

By Position: Statistically Non-Significant; Chi-square F<.143 df=2 n=1233

Q11(c) To what extent is the addition of other domains to the present TTAS instrument appropriate for appraisal of master teachers?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	39.2%	52.0%	52.8%	35.0%
Inadequately, Not at All	48.2%	48.0%	47.2%	65.0%
No Opinion	12.6%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1261

Q11(d) To what extent is the use of a different instrument and procedure appropriate for appraisal of master teachers?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	49.5%	56.2%	65.2%	62.5%
Inadequately, Not at All	31.9%	43.8%	34.8%	37.5%
No Opinion	18.6%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.045 df=2 n=1172

Somewhat related to the overall question concerning the relationship of the career ladder to the TTAS, the response percentages concerning appraising master teachers with the TTAS were neutral with superintendents moderately against this use. There was no doubt that respondents across all professional positions were strongly against using a different scoring procedure for the TTAS for master teacher purposes. The addition of other domains received neutral opinions from the administrators with teachers moderately opposed to this idea. Perhaps the strongest statement that can be made was the moderate support of teachers and superintendents for the use of both a different instrument and procedure for appraisal of master teachers. Berliner (1988), after examining the findings from a number of studies, reached the following conclusion:

> Instrumentation that is perfectly sensible to use with novices may not be suitable for proficient and expert teachers who have found their own ways of accomplishing the tasks of teaching. These teachers are more intuitive and holistic in their ways of teaching, and their classroom behavior may look different than that of novices or advanced beginners.

On-Site Interviews:

An additional question posed in the interviews that was not directly addressed in the questionnaire was "Is there potential for the use of the TTAS to identify the differences between master and regular teachers?" Overall, 70 percent felt that there was potential, but some of the teachers commented that the TTAS does not differentiate between teaching styles which might reflect differences between master and regular teachers. The four questions related to this issue from the questionnaire, when posed to those interviewees, received slightly more support than in the survey. The question concerning different scoring procedures [Q11(b)] had much the same response pattern as from the questionnaire. Those interviewed evidenced significantly more support for using the present instrument and, predictably, less support for the need for an entirely new procedure or instrument. Opinions concerning the need for new domains were equally split, much as in the questionnaire.

Open-Ended Comments:

(Principal) The goal of the evaluation should be to improve instruction. If the career ladder is tied to it, it creates major conflicts. Teachers will focus on the money instead of improvement.

(Principal) Do not tie TTAS to career ladder decisions, use it as a tool to improve classroom effectiveness.

(Superintendent) In its present form, TTAS leads to too much competition among teachers for career ladder dollars and is negative for students.

(Teacher) Raise salaries to a professional level and return integrity and professionalism to teaching.

(Teacher) Merit pay belongs to those who deserve it and not as a means to a quota system in order to use available money.

PEIMS Data:

Without repeating the data presented in Chapter 3, the information from PEIMS indicates some external influence over scores. One postulated influence is the close relationship of the TTAS to the career ladder. Please refer to the third chapter for a discussion of these data.

Other Uses of the TTAS

Contract Renewals and Campus Assignments. Another aspect of the use of the TTAS centers on contract renewal and campus assignment decisions.

Survey Questions:

Q16(a) To what extent should the TTAS be linked to contract renewal or termination?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	50.9%	68.8%	57.7%	36.8%
Inadequately, Not at All	46.8%	31.2%	42.3%	63.2%
No Opinion	2.3%	n/a	n/a	n/a

By Position: Statistically **Significant**; Chi-square F<.001 df=2 n=1407

Q16(c) To what extent should the TTAS be linked to Campus assignments?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	27.0%	33.5%	29.9%	24.0%
Inadequately, Not at All	67.5%	66.5%	70.1%	76.0%
No Opinion	5.5%	n/a	n/a	n/a

By Position: Statistically *Significant*; Chi-square F<.003 df=2 n=1360

Perhaps predictably, principals viewed the TTAS as a tool to help remove ineffective teachers from the school. Teachers, on the other hand, were not supportive of this position. This moderate, but opposite support with superintendents in a neutral position, indicates the expected difference between managers and staff. Superintendents were, apparently, caught in the middle between these two very distinct viewpoints. It is apparent that the TTAS was viewed as more than just a teacher improvement mechanism by the principals. There is no question that all respondents viewed the TTAS as inappropriate for use in making campus assignments.

On-Site Interviews:

Following the pattern where those interviewed were more positive in their perceptions of the appropriateness of many uses of the TTAS, almost 80 percent of those interviewed supported the concept of using the TTAS for contract renewals, while only 35 percent agreed that campus assignments should be so linked. It should be remembered that responses from an interview where identification of an individual is possible and responses from confidential questionnaires will be expected to differ.

Open-Ended Comments:

(Superintendent) If teachers cannot be outstanding in a 45 minute presentation with 2 weeks advance notice, their certificates should be removed.

Discussion and Recommendations

From the mixed responses to the question about linking the TTAS with the career ladder, it is difficult to make statements about the general relationship of the two from the information contained in this survey. With regard to use of the TTAS for appraising master teachers, it is evident that using different scoring procedures to accomplish this is not supported nor, generally, is the addition of other domains. There is moderate support for the use of a different instrument and procedure for master teacher appraisal. It is obvious that there is no support for the results from the TTAS to be used for campus assignments.

Perhaps the question of most interest is the one concerning contract renewal and termination. Principals are moderately in support of this use of the TTAS, teachers are moderately against it, and superintendents' opinions are between these two groups. According to Webster (1988), "negligent evaluation" is one of the ways in which employees use evaluations to sue their employers. Common personnel actions include issues related to promotion, discharge, lay-offs, and merit pay (Goddard, 1989). According to Weiss (1988), employers have only two options in response to the concerns about the legal side of performance appraisals: (1) do not publish any performance appraisal policies, or (2) if such policies are published, be sure to administer them fairly. As evident from data presented earlier, the majority of teachers appraised in Texas are rated highly, leading to little ability to discriminate among teachers based on their appraisal scores. Given the apparent evidence that the TTAS system, as practiced and regardless of concept or design, is influenced by factors external to the system, the denial of contracts or merit pay may be vulnerable to litigation.

Recommendation 4.1. Implement the recommendations contained in the third chapter.

These recommendations will disassociate the TTAS and the career ladder, at least to a certain extent, and add credibility to TTAS scores.

Recommendation 4.2. Do not use TTAS scores for campus assignments.

There is no support among administrators or teachers for this use of the TTAS and it is of questionable value. Assigning a majority of lower rated teachers to a single school is clearly not in the best interest of students. Assignment of highly rated teachers to problem campuses, especially without other compensation, may have an impact on teacher morale and retention.

Recommendation 4.3. Do not use the TTAS alone to make contract decisions.

The questionable validity of the TTAS, as evidenced by the current distribution of scores, precludes use of the TTAS alone to make employment decisions.

Recommendation 4.4. Create a voluntary appraisal program linked to incentive pay.

According to Brandt (1990), reasons for a voluntary system include:

- Some teachers, including very good ones, have legitimate reasons not to participate in teacher appraisal and incentive pay programs.
- Self-assessment becomes the first screening method in a voluntary program.
- With mandatory participation, the primary motivation is to avoid being uncooperative or unsatisfactory in job performance rather than the more positive motivation of attaining recognition, reward, or a sense of accomplishment.

Such a program could be tried in a sample of districts to determine the impact of such an approach. Other states, notably Tennessee, have used this type of program and could serve as a source of information concerning implementation and impact.

Results

Chapter 5: Impact of the TTAS

Normally an assessment or evaluation program is implemented because of a need, or a perception of a need, to change the process or product associated with the program. In the case of the TTAS, there was a need to ensure that assignment to the career ladder was based on demonstrated teaching competence. In addition, there was a perception that student performance in Texas was low to poor and that by improving teachers' skills, student outcomes would improve.

As part of the general program to ensure teacher quality, Texas has implemented several teacher testing programs. The first is the Texas Examination for the Certification of Administrators and Teachers (TECAT) designed to assure teachers' basic skills in reading and writing for teachers certified prior to 1986. The second is the Examination for the Certification of Educators in Texas (ExCET) for new teachers entering the profession, designed to provide an assurance of both subject matter and pedagogical knowledge. Finally, the Texas Master Teacher Examination (TMTE) has been developed as part of the process of identifying master teachers in the state. The TTAS is another method to improve and ensure quality teaching, but through observation of the process of instruction as opposed to performance on a test. The various teacher tests are, of course, point-in-time measures that the teacher has to pass once while the TTAS provides longitudinal monitoring of teaching behaviors. The first four questions in this series contain three that directly address the issue of overall impact of the TTAS on teaching and student performance. The fourth examines impact of the TTAS on the teacher.

Survey Questions:

Q8(b) In your district, does the TTAS improve classroom teaching methods?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	65.8%	79.8%	74.4%	53.5%
Inadequately, Not at All	33.1%	20.2%	25.6%	46.5%
No Opinion	1.1%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.010 df=2 n=1422

Q8(c) In your district, does the TTAS improve unsatisfactory teaching skills?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	55.6%	65.5%	64.5%	46.7%
Inadequately, Not at All	42.3%	34.5%	35.5%	53.3%
No Opinion	2.1%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1408

Q8(d) In your district, does the TTAS increase student achievement?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	43.6%	58.0%	48.7%	35.6%
Inadequately, Not at All	51.2%	42.0%	51.3%	64.4%
No Opinion	5.2%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.001 df=2 n=1367

Q8(i) In your district, does the TTAS improve the morale of teachers?

	Overall	Principals	Superintendents	Teachers
Completely or Adequately	11.8%	16.3%	8.7%	9.8%
Inadequately, Not at All	86.8%	83.7%	91.3%	90.2%
No Opinion	1.4%	n/a	n/a	n/a

By Position: Statistically Significant; Chi-square F<.002 df=2 n=1420

Although there was strong support that the TTAS has improved classroom teaching methods, there was only moderate support from the administrators that the TTAS has improved unsatisfactory teaching skills. One possible interpretation is that for most teachers, some improvement can be attributed to the TTAS, but for the less than satisfactory teacher, the impact may be minimal. It is possible that the improvement noted might be because the TTAS has increased awareness of research findings concerning effective teaching practices, rather than because it has operated as a vehicle for staff development in improving these techniques.

Administrators are neutral in opinion concerning the impact of the TTAS on student achievement, while teachers are moderately against this proposition. One thing of which there is no dispute, the TTAS has absolutely not helped teacher morale - the strength of the response to the question whether the TTAS has improved morale, in combination with the comments received on the survey and in interviews, results in the unavoidable conclusion that the TTAS has had a negative impact on teacher morale.

Literature clearly indicates that the consistency of a system is especially important when linked to merit pay because merit pay systems succeed only when employees perceive them as being fair and reliably administered. Anything less leads to a significant decline in morale (Thayer, 1987). Although the benefits of an effective appraisal system include improved overall morale as well as a dynamic and progressive work environment, an unmonitored system can lead to a failed system (Lee, 1989). Brandt (1990) indicates that negative effects of a poorly implemented program include lower morale and self-esteem, and unhealthy and cutthroat competition.

On-site:

While the site visits supported the ability of the TTAS to improve classroom teaching methods and unsatisfactory teaching skills, and to identify ineffective teaching practices, there was no consensus regarding the ability of the TTAS to increase student achievement.

Open-Ended Comments:

(Superintendent) The TTAS has been a positive step in the right direction to help teachers and administrators to speak a common language with regard to teacher appraisal.

(Superintendent) I have seen much improvement in teaching during the five years we've had TTAS. Teachers who study the instrument can and do learn to improve their teaching methods. They also learn how to give a "dog and pony show" in the presence of the evaluator. I'm not sure we will ever overcome teacher resentment of TTAS in general. People do not like to be made to do some things.

(Principal) TTAS has been the best thing that ever happened. It forces administrators to observe the teachers and it helps teachers become more effective and organized.

(Principal) I feel that the current system stifles teaching as well as learning. It also discourages teachers from sharing excellent lessons.

(Principal) As long as observations remain scheduled, some teachers will play the game very well. As a result, TTAS will not improve instruction.

(Teacher) The present system is demoralizing. Teachers who have exceeded expectations and are clearly outstanding should not be subjected to constant appraisals. Appraisers should not be under pressure from districts to keep scores down. This practice causes competition among teachers and does not result in better teaching practices. Teachers are not the enemy of the educational system.

(Teacher) From experience, I know that greater awareness of correct teaching techniques has come from instruction of those expectations, and not from an appraisal. The appraisal appears to me to have little, if anything, to do with improving teaching.

(Teacher) I do not think TTAS has helped the teaching profession. It has hurt teacher morale.

Outcome Data:

Outcome data are available over the time period in which the TTAS has been used across the state. While it is impossible to directly attribute gains or losses in student outcomes to one particular influence, relationships may be evident when examining these data. If, in fact, teacher skills are being improved through the use of the TTAS, then, lacking some drastic external influence, student scores should be improving.

Obviously, there are many potential influences on student test scores besides teacher skills. Two critical assumptions must be made explicit before even casual much less causal relationships are inferred. The first is that the TTAS leads to enhanced teacher knowledge and performance. The second is that increased teacher skills and more effective instruction lead to increased student outcomes. The efficacy of this relationship is, however, one of the assumptions underlying the various efforts to enhance the instructional capabilities of teachers.

TEAMS Scores. The Texas Educational Assessment of Minimum Skills (TEAMS) was the basic skills test given to all students in grades 3-11 at every odd-numbered grade level from 1986 through 1990. Since TEAMS scores by campus and district have been published in local newspapers, this test has become a powerful driver of instructional focus. Given an increased familiarity with test objectives over time coupled with a great emphasis on the importance of test outcomes, it is to be expected that with absolutely no changes in resources, student abilities, teacher skills, or other factors commonly associated with enhanced performance, scores will rise over time. However, it is also true that if any test scores are subject to change with improved instruction, basic skills performance should be among those most likely to be influenced. Table 9 presents TEAMS scores from 1986-1990.

TABLE 9
TEAMS Scale Scores From 1985-1989*

Mathematics

Grade Level	1985	1986	1987	1988	1989	1989-1985
Exit Level	740	761	769	774	781	+41
9th	781	788	798	798	801	+20
7th	787	808	831	835	835	+48
5th	783	803	819	821	823	+40
3rd	793	827	847	848	849	+56

English Language Arts(Exit) / Reading(other grades)

Grade Level	1985	1986	1987	1988	1989	1989-1985
Exit Level	775	785	785	794	792	+17
9th	783	784	788	801	795	+12
7th	772	784	790	795	802	+30
5th	790	792	808	802	803	+13
3rd	772	793	809	819	815	+43

*Source: Texas Education Agency, Division of Student Assessment:
Scale Scores range up to 999

Although the overall gain in scale scores shows a cumulative increase in students' scores over time, careful examination of the tables indicates that the majority of the increases occur from year one to year two of the testing program when students and teachers are rapidly becoming familiar with the test objectives. Further, it should also be noted that the increase in scores must be examined relative to the possible range of the scale scores being used in the testing program. Although the lowest scale score varies by year, subject area, and grade level, the possible range is approximately 400 to 500 points. The largest gains represent a very small percentage increase against the possible score range, and again, the majority of the increase occurred from the first to second year.

College Entrance Examination Scores. Given a relative dearth of student outcome data at the state level, information on college admissions tests is presented here as another, but restricted, view of student performance. Although the Scholastic Aptitude Test (SAT) and American College Testing Program (ACT) are not achievement tests, they do provide, when used in conjunction with other information such as grades and class standing, some measure of a student's preparation for college work. These tests do, of course, contain a great deal of academic material. Students who take such examinations are self-selected in the sense that they are generally those wishing to attend college and are not representative of the entire student population in Texas. The percentages of students taking these tests in Texas did not change dramatically over the time period included in Table 10.

TABLE 10
TEXAS SAT and ACT Scale Scores*

Scholastic Aptitude Test

Test	1985	1986	1987	1988	1989	1989-1985
Math	459	458	459	462	462	+3
Verbal	419	419	416	417	415	-4

American College Testing Program

Test	1985	1986	1987	1988	1989	1989-1985
Composite (includes all subject areas)	16.5	17.1	17.3	17.6	17.5	+1.0

*Source: CEEB and ACT

As shown in the table, there is a modest gain in ACT scores from 1985 to 1989. Over the same time period, the average composite score for the nation remained relatively constant. The average score for Texas for this test remains approximately one point below the national average. On the SAT, there is a gain in math scores and an equivalent loss in verbal scores in Texas as compared to a one point gain in math and a four point loss in verbal scores for the nation. These data do not argue for a significant increase in student performance over this time period as measured by these tests. Again, it is emphasized that college entrance examinations are not achievement tests and are not representative of all students in the state.

Norm-Referenced Test Data. Although norm-referenced data (NRT) were available only from selected districts for this evaluation, these data are presented as an additional reference from which to judge student performance. Results for fifth graders were selected for examination to contrast the results with TEAMS testing at this grade level and because it is one of the more stable measurement points of achievement testing.

TABLE 11
Data From Selected Districts

Fifth Grade Achievement Test Scores (Total Reading or Reading Comprehension National Percentiles)

District	1985	1986	1987	1988	1989	TEST
A	32[1]	39	39	n/a	41	ITBS
B	50	52	43	45	47	MAT6
C	51	53	51	50	53	ITBS
E	48	52	53	53	53	ITBS
F	old test		56	59	58	MAT6

[1] - Projected percentile from earlier norms

Again, there is little reason to suspect an overall increase in student achievement as measured by these achievement tests. These data were obtained from some of the larger school districts in the state and should be fairly representative of a large proportion of students in Texas. Although there were some societal and educational policy changes over this period of time, the large numbers of students involved in these testing programs ensure relative stability of the results.

Discussion and Recommendations:

Although administrators agree that improvement of classroom teaching methods can be attributed to the TTAS, only half of the teachers agree with this statement. There is less support for the contention that the TTAS has improved unsatisfactory teaching skills. In some respects this would seem counterintuitive. If the TTAS should do anything, according to earlier findings, it should improve unsatisfactory basic teaching skills. This finding might be interpreted as agreement that the TTAS exerts a generally positive influence on teaching, but less support that the TTAS improves specific skill areas.

There is no support from the survey that student achievement has been positively affected by the TTAS and little support from student performance data that the TTAS has been causal in gains. It is difficult to separate the impact of the TTAS on TEAMS scores because they were contemporaneous in implementation. There is almost always a large gain in student scores in the second year of a new testing program and it might be that the TTAS also had a substantial impact during this time as well.

This issue can be further explored in the future, with the implementation of a new testing program, the Texas Assessment of Academic Skills (TAAS) in 1990. Student performance was significantly depressed, in terms of percent passing all tests taken, in the first year of the TAAS. After testing in the fall of 1991, it is expected that an increase in scores similar to that seen with TEAMS - but smaller in magnitude because of a difference in measurement approaches - will be evident. If this is true, then the gain seen in the first years of TEAMS will be difficult to attribute to the TTAS. Other test scores, of course, did not show the dramatic increases seen with TEAMS scores. From the available data, if the TTAS was meant to increase teacher ability which would then increase student performance, the goal was not met.

Recommendation 5.1. Create two comparison groups of districts and campuses by continuing the TTAS in its present form in one group and replacing it with "walk-through" visits on a regular basis in the other group. Examine score trends for overall student performance on a variety of measures.

If the TTAS is continuing to make an improvement in teaching skills, then another evaluation of that impact using a similar questionnaire will substantiate this premise. Changes in student performance can also be used to determine more clearly the impact of the TTAS through teacher performance on student outcomes.

Recommendation 5.2. If changes are made to the TTAS or the career ladder, then a reassessment similar to this evaluation should be repeated at some point in time after implementation of the changes.

Continued monitoring on a periodic basis is recommended to ascertain usefulness and quality of the program, however changed.

Results

Chapter 6: Career Ladder

Clearly teachers and administrators closely link the career ladder and the TTAS, even though each system could exist independently of the other. The following question was included in the survey to explore this link.

Survey Question Q19:

OPINION PERCENTAGES WITHIN PROFESSIONAL CATEGORY

Options for the Career Ladder	Overall	Principals	Superintendents	Teachers
Abolish; Provide Monetary Incentives to High-Performing Campuses	11.2%	12.7%	14.1%	9.4%
Abolish	18.5%	22.3%	27.9%	12.3%
Abolish;Increase Salaries	41.8%	44.9%	44.9%	36.7%
Retain, as is	2.3%	1.9%	1.1%	3.1%
Retain, and Fully Fund	21.3%	14.6%	8.0%	31.9%
None of the Listed Options	5.0%	3.6%	4.0%	6.8%

PEIMS Data:

Paralleling the earlier discussion where total TTAS scores were examined by various teacher demographics, the following table presents the percentage of teachers at varying levels of the career ladder.

TABLE 12
Career Ladder Assignment by Teacher Demographics

Percentage of Teachers by Demographic and Career Ladder

Career Ladder	Sex		Ethnicity*		
Classification	Male	Female	Black	Hispanic	White
1	46.9	40.4	29.5	40.7	43.2
2	40.3	40.3	47.7	41.8	39.3
3	12.8	19.3	22.9	17.5	17.5
Totals	100.0	100.0	100.0	100.0	100.0

Similar to the findings presented in Table 7, where males received lower TTAS scores than females, the same relationship is evident in Table 12 where a higher percentage of females are on higher steps of the career

BEST COPY AVAILABLE

ladder. What is considerably different in this table is the relationship of the ethnic categories to career ladder placement. In the earlier table, the distribution of TTAS scores was very similar for the various ethnic groups. This is not true when examining career ladder distributions. There is a smaller percentage of Black teachers on the lowest level as compared to Hispanic and White teachers and a correspondingly higher percentage on the upper level. The Hispanic and White teachers are similarly distributed across the levels. It is not immediately apparent what the causal relationship might be, and further exploration of this pattern requires investigation that is beyond the scope of this study.

Although not presented, the relationships between career ladder placement and years of experience and between career ladder placement and degrees held reflect the patterns presented earlier between these variables and TTAS scores. The relationship is that more years of experience are associated with higher placement on the career ladder, with no change in this effect after 10 years of experience. There is little difference in career ladder level between the master's and Ph.D. with both having an advantage over the bachelor's degree.

Another issue related to the career ladder and the TTAS is the ability of the TTAS to distinguish between experienced and beginning teachers. The PEIMS data described earlier indicated that the overall mean score for teachers in the spring of 1990 was 157.4. During this same time, the mean score for teachers on career ladder level I was 151.7, on level II the mean score was 158.1, and for level III it was 167.7. Because career ladder assignment is confounded with TTAS scores in that high ratings from the TTAS are necessary to advance and remain on the career ladder, it is to be expected that mean scores on the TTAS will increase as career ladder levels increase. This pattern is evident, but the fact that most teachers, regardless of level, tend to have high TTAS scores argues for limitations in the range of scores at all levels. Any addition of domains to the TTAS in order to increase the ability of the instrument to discriminate higher levels of teaching will not do so if the overall average scores simply increase proportionally to the allowable maximum score.

Demographics of Districts. To parallel the TTAS analyses, the relationship between career ladder assignment and certain demographic characteristics of districts was examined using the agency's ANALYZE tool. Although there was a working hypothesis of a relationship between wealth per student of the district and number of teachers on higher steps of the career ladder, this was not evident from the data.

TABLE 13
Percent of Teachers on Career Ladder Steps (Wealth Categories)

	Percent of Teachers		
Wealth Category	Level One	Level Two	Level Three
Under $82,513	47.8	39.0	13.1
$ 82,513 to $ 99,523	43.3	40.2	16.5
$ 99,524 to $115,679	47.6	43.4	9.1
$115,680 to $133,839	42.0	42.8	15.2
$133,840 to $157,103	41.9	39.9	18.2
$157,104 to $181,840	45.6	42.1	12.3
$181,841 to $220,234	41.8	37.3	20.8
$220,235 to $287,975	34.9	43.4	21.8
$287,976 to $443,845	38.1	37.0	24.9
Over $443,845	42.5	42.3	15.2

When examining the next table, however, there is a relationship between district size as measured by average daily attendance (ADA) and between district type and percent of teachers on the career ladder.

TABLE 14
Percent of Teachers on Career Ladder Steps (District Size and Type)

	Percent of Teachers		
District Size(ADA)	Level One	Level Two	Level Three
Over 50,000	30.9	41.9	27.2
25,000 to 49,999	40.9	32.3	26.8
10,000 to 24,999	43.9	37.6	18.5
5,000 to 9,999	41.1	42.4	16.5
3,000 to 4,999	40.9	45.8	13.2
1,600 to 2,999	45.3	43.8	10.8
1,000 to 1,599	45.9	45.4	8.7
500 to 999	48.0	45.8	6.2
Under 500	53.1	40.6	6.2

	Percent of Teachers		
District Type	Level One	Level Two	Level Three
Major Urban	33.4	39.2	37.4
Major Suburban	42.0	34.7	23.4
Other Central City	42.7	40.3	17.0
Other CC Suburban	42.5	44.9	12.6
Independent Town	40.3	45.1	14.6
Non-Metro Fast Growing	50.2	40.2	9.6
Non-Metro Stable	45.2	45.8	9.1
Rural	51.1	42.6	6.2

There is a substantial relationship between both district size and district type and percentage of teachers on career ladder levels one and three. While several causal explanations have been proposed concerning this relationship, including the more likely presence, in larger districts, of teacher organizations and legal actions concerning career ladder placement no clear explanation was immediately apparent in this study, which of course concentrated on the TTAS. In addition, the following district characteristics were not systematically associated with higher percentages of teachers on career ladder level three.

The following district demographics were not associated with career ladder placement:
- Wealth
- Total Tax Effort
- Operating Cost/RADA
- Percent Passing All Tests Taken on TEAMS
- Percent Minority Students
- Percent Minority Teachers
- Percent Low Income Students

The following district demographics were associated with career ladder placement but are confounded with the measurement:
- Average Teacher Experience (marginal)
- Average Teacher Salary
- Percent Teachers with Advanced Degrees

Although no explanation is readily available, career ladder placement was associated with the price differential index (PDI) of the district and the Educational Service Center (ESC) Region. Percentages by region ranged from a high of 28.3 in Region XX to a low of 3.3 in Region IV.

While there may be some random associations present in these data, there are also effects which are consistent over years and seem to have been influenced by external forces.

Discussion and Recommendations:

Interestingly, opinions did not vary according to the level of the career ladder to which the teacher was assigned, nor to their years of experience. Support for maintaining the career ladder as it currently exists is minimal at best. Administrators clearly favor abolishing the career ladder with about one quarter indicating it should go, even with no alternative. Not surprisingly, almost half of the administrators and slightly over one-third of the teachers want to see the career ladder abolished and replaced with increased salaries. It was somewhat surprising to see that almost one-third of the teachers wanted to retain the ladder, but to fully fund it, an option not espoused by the administrators.

Recommendation: Given a certain limit on funding, a compromise position which considers the results of the survey would be to abolish the career ladder, and determine alternative methods of allocating funds.

One possible alternative includes use of the Academic Excellence Indicator System (AEIS) to reward campuses across all types of schools which exhibit exceptionally high levels of performance. The AEIS provides an opportunity to show high levels of performance relative to similar schools.

OVERALL SUMMARY

An evaluation was conducted using a variety of quantitative and qualitative data addressing the TTAS and, to a certain extent, the career ladder. The analyses generally indicated that the TTAS as implemented during the previous two school years did not result in discrimination among various performance levels of teachers. This conclusion was based on the finding that use of the TTAS resulted in repetitive scores, patterning of score frequencies, and use of only two of the five available summary categories. Although most respondents agreed that the TTAS has potential as a teacher appraisal system, it is not being used in a manner consistent with good practice.

Part of the difficulty with the lack of discrimination in TTAS scores may be due to the relationship of the TTAS to the career ladder. There was no support from the results of this study for keeping the career ladder as currently implemented in districts. As detailed in the report, recommendations in this area are based on the assumption that the TTAS will be disassociated from the career ladder. After this has been accomplished, several approaches to improving the TTAS have potential, such as restricting the use of the TTAS to beginning teachers, including objective measures of student performance, and making participation in the system a voluntary one.

Included in Appendix F are comments from the Assistant Commissioner for Professional Development at the Texas Education Agency. The Assistant Commissioner is responsible for oversight of the TTAS among other responsibilities. Comments include reflections on this evaluation and steps taken by the Division of Management Assistance and Personnel Development to correct some of the difficulties evident with the TTAS.

As part of the charge to the outside contractor, an independent report on the TTAS as practiced in Texas was produced and is available as a separate document from the agency. On the following pages, there is a recapitulation of the recommendations that resulted from the findings of this evaluation.

Summary of Recommendations

Recommendation 1.1.	Investigate the relationship of student outcomes, especially as included in the Academic Excellence Indicator System, to individual teachers and to campus-level aggregations of appraisal scores.
Recommendation 1.2.	Investigate the potential for adding new domains or criteria to the TTAS which would assess the teachers' abilities to teach critical thinking skills and to teach students of different ability levels.
Recommendation 2.1.	Establish a pilot program in selected districts under careful controls to examine awarding EQs at the indicator level.

Recommendation 2.2.	Establish a pilot program in selected districts under careful controls to examine documentation of EQs. Restriction, enhancement, or removal of requirements for documentation should be examined.
Recommendation 2.3.	After the first two years of teaching, eliminate annual evaluations. Appraise every third year for a minimum of four observation periods, only one of which would be scheduled.
Recommendation 3.1.	Disassociate the TTAS from the career ladder.
Recommendation 3.2.	Evaluate appraisers based on a distribution of scores awarded over time, especially as compared to statewide information. Publish means and distributions of campus, district, and statewide TTAS scores.
Recommendation 3.3.	Continue the educational process for appraisers, teachers, and the general public concerning the meaning and utility of the various rating categories. Universities associated with teacher training need to strengthen curriculum in this area.
Recommendation 3.4.	Investigate the use of a cadre of "professional" appraisers, perhaps based at the education service centers.
Recommendation 4.2.	Do not use TTAS scores for campus assignments.
Recommendation 4.3.	Do not use the TTAS alone to make contract decisions.
Recommendation 4.4.	Create a voluntary appraisal program linked to incentive pay.
Recommendation 5.1.	Create two comparison groups of districts and campuses by continuing the TTAS in its present form in one group and replacing it with "walk-through" visits on a regular basis in the other group. Examine score trends for overall student performance on a variety of measures.
Recommendation 5.2.	If changes are made to the TTAS or the career ladder, then a reassessment similar to this evaluation should be repeated at some point in time after implementation of the changes.
Recommendation:	Given a certain limit on funding, a compromise position which considers the results of the survey would be to abolish the career ladder, and determine alternative methods of allocating funds.

APPENDIX A
Historical Review and SBOE Rules

Texas Teacher Appraisal System (TTAS)
Teacher Orientation

- Each continuing teacher must receive a copy of the most current Teacher Orientation Manual or supplement thereto and an orientation to any changes in the TTAS should be provided.

- Each new teacher must receive a copy of the most current TTAS Teacher Orientation Manual and an orientation to the TTAS prior to the teacher's first observation.

SUGGESTED FORMAT FOR NEW TEACHER ORIENTATION

1. Communicate the following background information regarding TTAS:

    - Historical Development
    - TTAS Assumptions
    - TTAS Development Process

2. Review state and local requirements:

    - SBOE Rules for teacher appraisal
    - District Policy/Procedures

3. Discuss each of the following aspects of the TTAS and review the forms used in the appraisal process:

    - Teacher Assessment of Instructional Goals and Outcomes (TAIGO)
    - Professional Growth Plan (PGP)
    - Cumulative Data
    - Exceptional Quality

4. Review the TTAS Instrument using the "Introduction to the Appraisal System, Domains I-V" videotape.

5. Explain scoring procedures:

    - Observation Record/Evaluation Record Form (OR/ER)
    - Appraisal Record Form (AR)
    - Score Conversion Chart

6. Discuss observation procedures:

    - Calendar (dissemination of time frame for appraisal periods 1, 2, and single)
    - Waiver of 45-minute observation requirement
    - Local policies (in regard to scheduling)
    - Local staff development program addressing appraisal
    - Pre-observation conference (if required by local policy)
    - Post-observation conferences
    - Summative conferences

Historical Development

House Bill 72, enacted by the 68th Legislature, second called session, directed the State Board of Education to adopt an appraisal process and criteria with which to appraise the performance of teachers for career ladder purposes. The Legislature mandated that the criteria be based upon observable, job-related behaviors. The State Board of Education was further directed to provide: (1) at least two appraisals during each of the two appraisal periods within the regular school year; (2) a uniform training program for appraisers of teacher performance, including uniform appraiser certification standards; and (3) inclusion of teacher self-appraisal in the appraisal process. In response to this mandate, the State Board of Education adopted a timeline for development of the teacher appraisal system in January of 1985. To assist in this endeavor, a Teacher Appraisal Advisory Committee (TAAC) composed of teachers, administrators, regional education service center personnel, and representatives from institutions of higher education, was appointed by the State Board of Education.

Texas Education Agency staff conducted a review of literature on teaching effectiveness, surveyed other states where statewide appraisal systems had been implemented, and gathered information from 156 school districts in Texas regarding the teacher evaluation systems used in those districts. Subsequently, a job-relatedness survey, which had as items those behaviors identified in the previous processes, was conducted. Thirty thousand teachers, stratified by gender, race/ethnicity, teaching field, teaching assignment, and years of experience were asked to respond to items in terms of observability, importance, and frequency of use. Approximately 17,000 responses were received. A list of teaching behaviors to be considered for inclusion in the appraisal instrument was derived from this survey.

The State Board of Education Committee on Personnel reviewed materials, directed revisions, and authorized an expert review of the draft materials in September of 1985. Nationally recognized educators participating in the review of materials included Dr. John Goodlad of the University of Washington, Dr. Richard Manatt of Iowa State University, Dr. Lester Solomon of the Georgia State Department of Education, and Brigadier General Billy Bowles, a military evaluation and staff development expert.

In October of 1985, a pilot program was implemented in six school districts: Slaton, Santa Rosa, Seguin, Grandfalls-Royalty, Port Arthur, and New Boston. Ninety local appraisers in these districts were trained by Performance Assessment Systems, Inc. of Athens, Georgia to use the appraisal instrument and procedures. More than 1,400 teachers were appraised during the pilot study. Further revisions of the instrument and procedures were made during this training. Data regarding such factors as time requirements, attitudes, and usability of performance indicators were collected during this study.

On February 6 and 8 of 1986, comments and concerns of teachers, administrators, and professional organizations were given at a public hearing conducted by the State Board of Education. Revisions of the appraisal system were made as a result of the public hearings and of the findings of the pilot study.

In April and May of 1986 approximately 270 individuals were trained to deliver training to school district appraisers during the summer of 1986. Forty-six percent of these trainers were staff members from regional education service centers, fifty-two percent were school district personnel, and two percent were staff from institutes of higher education.

Approximately 13,000 appraisers received the initial 43-hour training during the summer of 1986. Course content included study of the statutory requirements and State Board of Education rule addressing the appraisal of certified personnel, the appraisal instrument and procedures, including scoring and practice in application of the appraisal procedures under simulated conditions. Each participant was tested on knowledge of the system and on proficiency in scoring videotaped segments of instruction. A standard of 70 percent proficiency was set by the State Board of Education.

The Texas Teacher Appraisal System was implemented on a statewide basis in the fall of 1986. The first appraisal period was designated as a practice period and actual summative appraisals began during the second appraisal period. Some changes (i.e., the removal of zeros from the observation record form and a modification of the requirements for professional growth plans) were made by the State Board of Education in January 1987. Also, standards were set for the rating of teacher performance for career ladder decisions.

In February of 1987 an advisory committee composed of teachers, administrators, professional organizations, and education service center appraisal trainers met in Austin to develop recommendations regarding areas for improvement of the appraisal system. The findings of that group were presented to the Committee for Personnel at the March meeting of the State Board of Education. The recommendations of the advisory committee as well as the analyses of data collected during the first appraisal period provided the bases for further refinement and modification of the TTAS. Refinements to date include a reduction in the number of criteria and the number of indicators, development of a whole instrument scoring procedure, simplification of the Teacher Assessment of Instructional Goals and Outcomes, and movement of exceptional quality scoring to the criterion level.

Training updates and proficiency checks were completed by all eligible appraisers during the summer of 1987, with modifications implemented on September 1, 1987.

In the summer of 1987, the Second Called Session of the 70th Legislature passed House Bill 173, which reflected changes in the state appraisal system. Amendments to 19 TAC, Chapter 149, Subchapter C, were made in response to this bill to include appraisal by grade level or department chairpersons, reduction in the number of required appraisals for eligible teachers, and evaluation of nondegreed teachers. Later that year the State Board of Education provided for local board approved leaves of absence for teachers without career ladder penalty.

Assumptions Underlying the Texas Teacher Appraisal System

The Texas Teacher Appraisal System developed to implement House Bill 72 is intended to encourage professional growth for both teachers and administrators and to improve instruction in the classrooms of Texas. In designing the system, efforts were made to build a practical, usable system that can be applied fairly to teachers in all subjects and grade levels. The teacher appraisal system does not attempt to assess all aspects, duties, and responsibilities of teaching. In part, this is because the legislative requirements provide a more narrow focus and, in part, because the state of the art of teacher evaluation is not advanced to an operational level in some areas. Instead, the system has been based upon existing classroom-based research on teaching, craft knowledge and experience.

The appraisal process has also been designed to include principles of sound evaluation and to reflect the best current practices in the field of teacher evaluation.

The teacher appraisal process emphasizes the professional growth of teachers, but because the teacher appraisal system was designed as a part of the career ladder system for teachers established by the Legislature, it also must serve the function of distinguishing among the performances across the state. In order to do so, the teacher appraisal process is designed to collect samples of valid information about teaching in a manner that can support reliable decisions about teacher performance.

Certain assumptions which serve as the foundation for the statewide system are stated below. Some of these assumptions are related directly to the statutory requirements, while others arise from educational research and experience.

View of Teaching: An appraisal system is grounded in particular views of teaching that are reflected in the instruments and processes associated with that appraisal system. The Texas teacher appraisal system is based upon the assumption that teaching is an intentional act which has as its goal student growth. Teaching includes, but is not limited to, "planning, delivering, evaluating, and reporting of student learning of the essential elements" as required by statute and regulation. Although classroom teaching is the primary focus of the performance appraisal, no single model of teaching is mandated by the statewide teacher appraisal system.

Generic Teaching Behavior: As required by legislation, all teachers will be appraised with the same criteria and in the same manner regardless of career ladder assignment. Although teaching varies across subject matter and grade level, some common teaching behaviors occur which can serve as areas of evaluation.

Quality of Performance: Many of the skills necessary to be a successful master teacher are not different in kind from those necessary to be a successful beginning teacher. The difference between these teachers will appear in the number of skills exhibited by the teachers at levels of quality that meet or exceed stated expectations. Most teachers will not exhibit all of the behaviors with a quality of performance which always exceeds expectations.

Professional Growth: The teacher appraisal process assumes that each teacher is capable of improving regardless of the level of expertise or years of experience.

Educational Goals: While student academic growth is the primary goal of the public school system in Texas, that goal is supplemented by other student development goals. As a consequence, appropriate classroom teaching behaviors may vary according to the learning goals and objectives of the lesson observed.

Accuracy of Appraisal: Appraisal of teaching performance requires classroom observation which, by its very nature, intervenes in the classroom. Therefore, the observers will be seeing the teacher exhibit what he or she can do, not necessarily what the teacher usually does when an appraiser is not in the classroom. More accurate appraisal will be gained by using multiple appraisers and multiple sources of information.

An Overview of the Texas Teacher Appraisal System

Part A. Classroom Teaching Performance Domains

Domains:
- I. Instructional Strategies
- II. Classroom Management and Organization
- III. Presentation of Subject Matter
- IV. Learning Environment

These domains will be evaluated by a minimum of two certified appraisers through formal classroom observation. If additional information related to the performance criteria has been documented and shared with the teacher, it may be considered by the teacher's supervisor. Appraisers will be trained and certified through the state training program. The results of the appraisal will be used for career ladder assignment and staff development and may be used for contract renewal considerations.

Part B. Professional Growth and Responsibilities Domain

Domain:
- V. Professional Growth and Responsibilities

This domain will be evaluated by the teacher's supervisor. Credit will be awarded for each indicator unless documentation on file supports the denial of credit. If the teacher's Domain V score as assigned by the teacher's supervisor is less than satisfactory, the other appraiser will review the documentation and assign a score for Domain V. Training will focus on due process procedures and documentation methods. The results will be used for career ladder assignment and staff development and may be used for contract renewal considerations.

Texas Teacher Appraisal System

Teacher Orientation

Changes in TTAS for continuing teachers
New teachers prior to first observation

Teacher Assessment of Instructional Goals and Outcomes (TAIGO)—Teacher
Complete Part 1 prior to end of 2nd six weeks

Observation 1—Teacher's Supervisor
- Pre-observation conference (optional)
- Complete at least one formal observation
- Complete Observation Record **(OR)**, provide copy to teacher within seven working days
- Conduct conference within ten working days

Observation 2—Other Appraiser(s)
- Pre-observation conference (optional)
- Complete at least one formal observation
- Complete Evaluation Record **(ER)**, provide copy to teacher within seven working days
- Conduct conference within ten working days if performance is less than meets expectations in one or more domains

Appraisal 1—Teacher Supervisor
- Review any additional information (informal observations/other documentation)
- Complete ER
- Develop professional growth plan if performance less than satisfactory in one or more domains (two appraisal teachers only)
- Conduct conference at teacher's or appraiser's request

Observation 3—Teacher's Supervisor
- Pre-observation conference (optional)
- Complete at least one formal observation
- Complete **OR**, provide copy to teacher within seven working days
- Conduct conference within ten working days (may be combined with the summative conference)

Observation 4—Other Appraiser(s)
- Pre-observation conference (optional)
- Complete at least one formal observation
- Complete ER, provide copy to teacher within seven working days
- Conduct conference within ten working days if performance is less than satisfactory in one or more domains (may be combined with the summative conference)

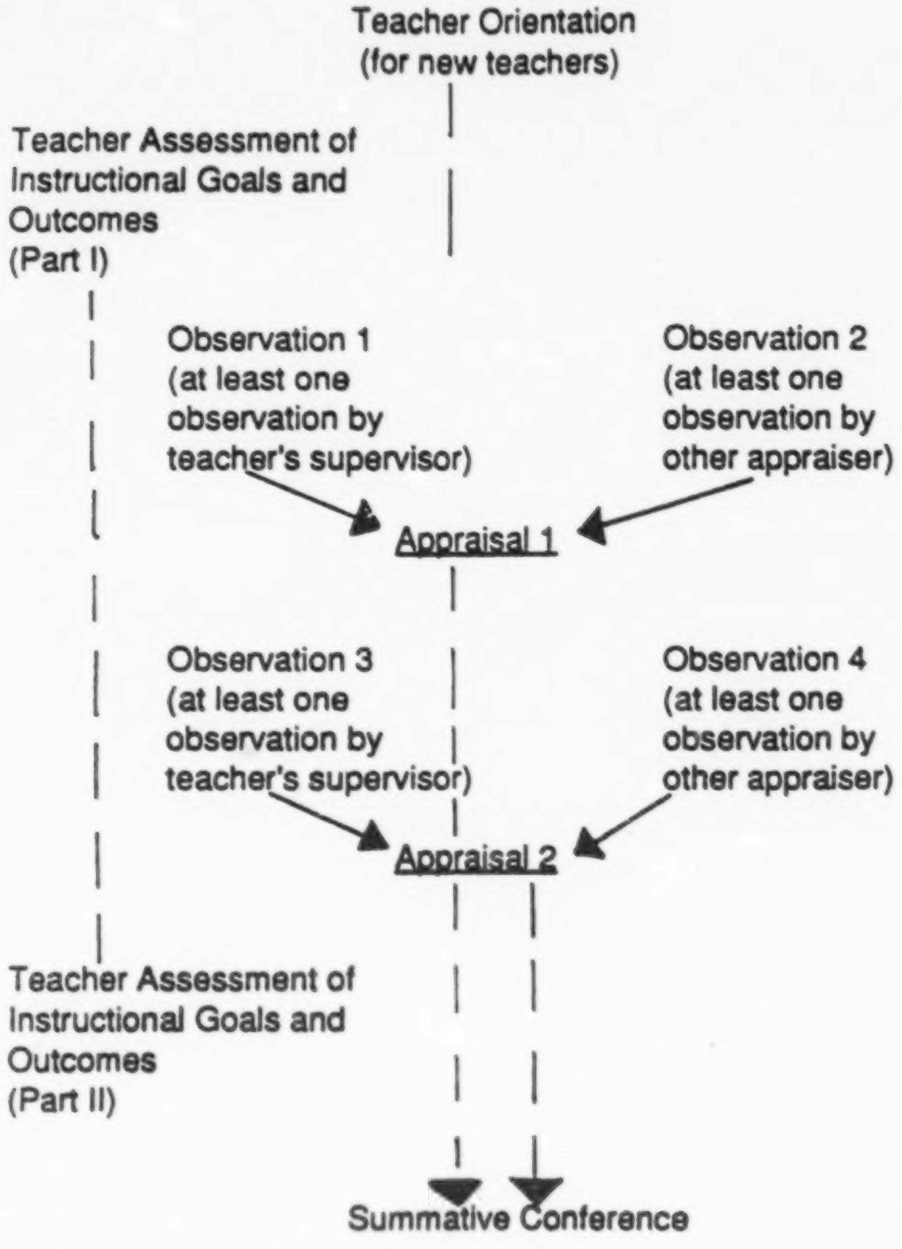
Appraisal 2—Teacher Supervisor
- Review any additional information (informal observations/other documentation)
- Complete ER
- If Domain V score is less than satisfactory, ask other appraiser(s) to review documentation for Domain V, score ER, provide copy to teacher and supervisor
- Complete Appraisal Record **(AR)**

Teacher Assessment of Instructional Goals and Outcomes (TAIGO)—Teacher
Complete Part II prior to summative conference

Summative Conference—Teacher Supervisor
- Review completed TAIGO
- Advise teacher of Domain and Overall Summary Performance score(s); provide copy of AR to teacher
- Discuss teacher's status for maintenance/advancement on career ladder
- Develop/modify a professional growth plan if Overall Summary Performance is less than satisfactory; at teacher's request if Overall Summary Performance meets expectations
- Make recommendations regarding need for improvement in any domain(s)

R-89/90

TEXAS TEACHER APPRAISAL SYSTEM


Teacher Orientation (for new teachers)
Teacher Assessment of Instructional Goals and Outcomes (Part I)
Observation 1 (at least one observation by teacher's supervisor)
Observation 2 (at least one observation by other appraiser)
Appraisal 1
Observation 3 (at least one observation by teacher's supervisor)
Observation 4 (at least one observation by other appraiser)
Appraisal 2
Teacher Assessment of Instructional Goals and Outcomes (Part II)
Summative Conference

July 1990

Current State Criteria for Advancement and Maintenance on Career Ladder

	(Fast Track)	(Regular Track)
	LEVEL TWO	
Year of Implementation:	1984-85	1984-85
Experience	Two Years	Three years
Training:	Masters degree (job-related)	Bachelors degree plus 9 sem./135 AAT hours additional training
Entry Performance:	Exceeds expectations in preceding year	Same
Maintenance Performance:	Meets expectations each year	Same
	LEVEL THREE	
Year of Implementation:	1987-88	1989-90
Experience:	Three years at level two	Five years at level two
Training:	3 sem./45 AAT hours additional training	6 sem./90 AAT hours additional training
Entry performance:	Clearly outstanding in 2 of last 3 years	Exceeds expectations in 3 of last 4 years
Maintenance performance;	Exceeds expectations 2 out of every 3 years	Same

STATE BOARD
OF
EDUCATION
RULE

R-89/90

SUBCHAPTER C. APPRAISAL OF CERTIFIED PERSONNEL

§149.41 General Provisions.

Statutory Citation

Texas Education Code, §13.302:

"(a) The State Board of Education shall adopt an appraisal process and criteria on which to appraise the performance of teachers for career ladder level assignment purposes. The criteria must be based on observable, job-related behavior, including teachers' implementation of discipline management procedures.

"(b) The board shall solicit and consider the advice of teachers in developing the appraisal process and performance criteria.

"(c) In developing the appraisal process, the board shall provide for using not fewer than two appraisers for each appraisal. One appraiser must be the teacher's supervisor and one must be a person as approved by the board of trustees. An appraiser who is a classroom teacher may not appraise the performance of another classroom teacher who teaches at the same school campus at which the appraiser teaches, unless it is impractical because of the number of campuses or unless the appraiser is the chairman of a department or grade level whose job description includes classroom observation responsibilities. The board also shall provide for a uniform training program and uniform certification standards for appraisers to be used throughout the state. The board shall include teacher self-appraisal in the process.

"(d) The State Board of Education shall develop or adopt and validate an assessment instrument which may be administered to teaching personnel for the purpose of evaluating the performance of those personnel in the jobs for which they were hired. The assessment process may:

(1) be administered by or under the supervision of the Central Education Agency on a statewide basis;

(2) provide opportunities for preparation and remediation;

(3) provide reasonable opportunities for individuals to attain an acceptable score set by the board;

(4) include provision for substitution of the results of alternative assessment instruments approved by the board;

84

(5) be an integral part of the appraisal process and shall not be considered as a substitute for any evaluation by observation as may otherwise be required; and

(6) be administered by local districts.

"(e) It is the intent of subsection (d) of this section that the assessment instrument described therein be used to assess specific skills primarily for the purpose of remediation and improvement.

"(f) Appraisal for teachers must be detailed by category of professional skill and characteristic and must provide for separate ratings per category. The appraisal process shall guarantee a conference between teacher and appraisers, and the conference shall be diagnostic and prescriptive with regard to remediation as needed in overall summary of performance by category and identify the required performance for advancement to the next level."

Texas Education Code, §13.303:

"(a) Each school district shall use the appraisal process and performance criteria developed by the board in appraising teachers for career ladder level assignment purposes.

"(b) The school district shall determine the number of appraisers used if the number is to exceed the minimum required.

"(c) Appraisal shall be done not fewer than:

(1) two times during each school year for probationary teachers and for teachers on level one of the career ladder; and

(2) once during each school year for teachers on levels two, three, and four of the career ladder whose performance, on the most recent appraisal, was evaluated as exceeding expectations or clearly outstanding. The performance of a teacher who, because of unusual circumstances, is appraised only once in a particular year shall be evaluated for career ladder purposes on the basis of a single appraisal."

Texas Education Code, §13.304:

"In appraisals of teacher performance for career ladder level assignment purposes, performance shall be evaluated in the same manner and under the same criteria regardless of level. Performance shall be evaluated as:

(1) unsatisfactory (if the teacher's performance is clearly not acceptable in some major area);

(2) below expectations (if the teacher's performance needs improvement in some major areas);

(3) satisfactory (if the teacher's performance meets expectations);

(4) exceeding expectations (if the teacher's performance excels in some major areas); or

(5) clearly outstanding."

Texas Education Code, §13.318:

"A teacher who directs extracurricular activities in addition to performing classroom teaching duties shall be appraised only on the basis of classroom teaching performance and not on performance in connection with the extracurricular activities."

Texas Education Code, §13.354, Appraisal of Administrators:

(a) The State Board of Education shall adopt an appraisal process and criteria on which to appraise the performance of school administrators. The criteria must be based on job-related performance.

(b) The board may solicit and consider the advice of teachers and administrators in developing the appraisal process and performance criteria.

(c) In developing the appraisal process, the board shall establish the qualifications required to serve as an appraiser and the number of appraisers to be used for each appraisal.

(d) Each school district shall use the appraisal process and performance criteria developed by the board in evaluating the performance of an administrator."

Rule

(a) Each teacher and administrator shall be appraised annually in the performance of his or her duties. This requirement shall apply to all certified personnel, as classified in the Texas Education Code, §16.056, Texas Public Education Compensation Plan, including the non-degreed teacher.

(b) The results of the appraisal of teachers shall be used for career ladder and staff development purposes and may be used for contract renewal considerations.

(c) For professional personnel exempted from the career ladder under §149.71 (b)(7) of this title (relating to Assignment to the Teacher Career Ladder), at least one evaluation conducted by one appraiser is required, using the evaluation instrument(s) adopted by the local board of trustees. Non-degreed teachers shall be evaluated once each year by two appraisers.

(1) Until the State Board of Education adopts an appraisal process for administrators, the evaluation instrument(s) adopted by the local board of trustees shall be used.

(2) Each appraiser shall be trained in the appropriate use of the instrument(s).

(d) Each school district shall appraise teachers using the state appraisal process and performance criteria developed and approved by the State Board of Education.

(e) As an exception to subsection (d) of this section, districts approved by the State Board of Education to conduct research projects for the state may be exempted from certain provisions of the state appraisal system. Approved research projects will address the use of student progress, a different approach to the identification of quality aspects of teacher performance, alternate procedures for scoring at the criterion level, or other issues of merit.

§149.42 Teacher Performance Criteria.

In the appraisal of teacher performance for career ladder assignments and for staff development, performance shall be appraised with the following performance criteria subsumed under five major areas henceforth called domains:

(1) Domain I. Instructional strategies.

(A) Criterion 1. Provides opportunities for students to participate actively and successfully.

(B) Criterion 2. Evaluates and provides feedback on student progress during instruction.

(2) Domain II. Classroom management and organization.

(A) Criterion 3. Organizes materials and students.

(B) Criterion 4. Maximizes amount of time available for instruction.

(C) Criterion 5. Manages student behavior.

(3) Domain III. Presentation of subject matter.

  (A) Criterion 6. Teaches for cognitive, affective and/or psychomotor learning.

  (B) Criterion 7. Uses effective communications skills.

(4) Domain IV. Learning environment.

  (A) Criterion 8. Uses strategies to motivate students for learning.

  (B) Criterion 9. Maintains supportive environment.

(5) Domain V. Professional growth and responsibilities.

  (A) Criterion 10. Plans for and engages in professional development.

  (B) Criterion 11. Interacts and communicates effectively with parents.

  (C) Criterion 12. Complies with policies, operating procedures, and requirements.

  (D) Criterion 13. Promotes and evaluates student growth.

§149.43 Teacher Appraisal Procedures.

(a) Appraiser qualifications.

  (1) The teacher appraisal process requires at least two appraisers for each appraisal period.

  (2) One appraiser must be the teacher's supervisor and, except for those serving as part-time principals or in other extraordinary circumstances approved by the commissioner of education, must hold administrator or supervisor certification.

  (3) Appraisers other than the teacher's supervisor must be approved by the local board of trustees, have a valid teaching certificate, and have at least two years of pre-kindergarten, kindergarten, elementary, or secondary classroom teaching experience.

(4) Classroom teachers may not appraise other classroom teachers on the same campus unless there is only one school campus in the district or unless the appraiser is the department or grade level chairman whose job description includes appraisal responsibilities. Teachers serving as appraisers must have an assignment on the career ladder at least as high as the teacher being appraised.

(5) Before conducting appraisals, each appraiser must receive instructional leadership training, must receive uniform appraiser training, and must reach the required standard of proficiency as established by the State Board of Education. Coursework equivalent to the required instructional leadership training may be substituted if completed within three years prior to receiving appraisal. Periodic recertification and appraisal training updates will be required for each appraiser.

(b) Teacher orientation. Each district will provide each teacher with an orientation to the teacher appraisal system prior to the teacher's first appraisal with the system. As early as possible, but no later than the beginning of the orientation, each teacher must receive a copy of the teacher appraisal orientation manual. The Central Education Agency will provide materials for additional training of teachers as part of the local district's inservice program.

(c) Appraisals, observations, and conferences.

(1) At least two appraisals are required each year for each probationary teacher, each career ladder level one teacher and each teacher on career ladder levels two, three, or four whose performance on the most recent overall summary performance score was less than exceeding expectations. Teachers on levels two, three, or four of the career ladder whose performance was exceeding expectations or clearly outstanding on the most recent overall summary performance score shall be appraised at least once each year.

(2) One appraisal consists of a minimum of two formal 45-minute observations, one by each appraiser.

(3) All appraisals must be conducted during the required days of instruction for students during one school year. Excluding the first two weeks of instruction, each local district shall establish a calendar which designates the time frame for the appraisal periods. For all teachers who are eligible for a

single appraisal within a district, that appraisal may be uniformly extended throughout the days of instruction at district discretion. This calendar shall be disseminated to all staff prior to the beginning of formal observations. Formal observations shall not be conducted on the last instructional day before any official school holiday or on any other days deemed inappropriate by the local board of trustees.

(4) For formal observations, teachers shall be observed teaching classes in field(s) and teaching assignments for which they are certified whenever possible.

(5) For the 1989-1990 school year, 50 percent of the formal observations must be scheduled by day and time of day, and 50 percent of the formal observations must be scheduled within a reasonable period of time designated by the local district and uniformly applied for all teachers.

For the 1990-1991 school year and upon the development and approval of instruments, processes, or procedures to be used for purposes of appraising levels three and four of the career ladder, all formal observations using the Texas Teacher Appraisal System shall be scheduled.

(6) Before the first observation of the teacher in any appraisal period, the requirement for consecutive minutes for unscheduled formal observations may be waived by mutual consent at the request of that teacher or the appraiser. Under such waiver, each unscheduled formal observation may be comprised of two to three instructional segments of not less than 15 minutes each. Such waiver should be considered only when the nature of the teaching assignment requires shorter instructional segments.

(7) Appraisers may not conduct formal observations simultaneously.

(8) After a formal observation each appraiser must complete a written record. The written record is not to be completed during the observation. A copy of the written record shall be given to the teacher within seven working days of the formal observation. If there are extenuating circumstances, the seven working day requirement may be extended to a maximum of 15 working days.

(9) During an appraisal period, the teacher's supervisor may continually evaluate and document performance specifically related to the performance criteria and the indicators subsumed under the criteria in §149.42 of this title (relating to Teacher Performance Criteria). If such documentation would influence the teacher's appraisal, the documentation must be shared in writing with the teacher within seven working days of the occurrence or, in unusual circumstances, the teacher supervisor's knowledge of the occurrence. This additional documentation shall be combined with, but shall not replace, the formal observation to determine credit for the criteria or indicators. Appraisers other than the teacher's supervisor shall have access to Domain V documentation only in the event that the teacher's total score for the year on Domain V determined by the teacher's supervisor is less than meets expectations as specified in §149.44 (b)(2) of this title (relating to Teacher Appraisal Instrument, Scoring Procedures, and Forms).

(10) Following each formal observation, an appraiser must conduct a post-observation conference with the teacher if the performance is judged less than meets expectations in one or more domains. Regardless of the teacher performance, each teacher supervisor must conduct a post-observation conference. Appraisers other than the teacher supervisor are encouraged to conduct post-conferences after all formal observations. Required post-conferences must be held within ten working days of the formal observation. If there are extenuating circumstances, the ten working day requirement may be extended to a maximum of 15 working days. At the conclusion of the first appraisal period, a conference will be held at the request of either the teacher or the appraiser.

(d) Teacher response and appeals.

(1) A teacher may submit a written response regarding the appraiser's written record within 10 working days of receipt of that record. If there are extenuating circumstances, the ten working day requirement may be extended to a maximum of 15 working days. Such written response shall become part of the appraisal record.

(2) Each local district shall adopt a written policy establishing a procedure for a teacher to present grievances and receive written response regarding the evaluation process. This policy shall be disseminated at the time of employment to each professional staff member and updated annually or as needed. The judgment of the commissioner of education shall not be substituted for that of the local district unless the district's decision was arbitrary and capricious or made in bad faith.

R-89/90

(e) Professional growth plan.

(1) For teachers receiving two appraisals, a professional growth plan must be developed or modified if any domain is judged less than meets expectations at the end of the first appraisal period. However, professional growth activities may be planned and implemented any time at the request of the teacher or the appraiser(s).

(2) A professional growth plan must be developed or modified for any teacher whose overall summary performance score is less than satisfactory. At the teacher's request, a professional growth plan must be developed or modified for any teacher whose overall summary performance score is satisfactory.

(3) The teacher's supervisor, in cooperation with the teacher, will develop the growth plan. Other appraisers, as appropriate, may participate in this process. Options for growth activities shall be provided and at least one option shall place no significant financial burden on either the teacher or the local district. Fulfillment of the provisions of a professional growth plan does not, in and of itself, serve as a guarantee of career ladder advancement.

(f) Teacher assessment of instructional goals and outcomes.

(1) No later than the end of the second six weeks, each teacher will complete the first part of the teacher assessment of instructional goals and outcomes and file a copy with the teacher's supervisor. Significant changes in assignment or class composition may result in the need to update the first part of the teacher's assessment.

(2) During the week preceding the summative conference, the teacher will complete the second part of the assessment of instructional goals and outcomes. A copy of the completed assessment will be attached to the appraisal record at the summative conference.

(g) Summative appraisal.

(1) Each teacher must receive a summative conference at the end of the teacher's last appraisal period. In this conference, the teacher's supervisor will review the completed assessment of instructional goals and outcomes, inform the teacher of the domain performance scores and the overall summary performance score for the year, review the teacher's status relating to requirements for advancement and/or maintenance on the teacher career ladder, make recommendations regarding domains needing improvement, and address a professional growth plan as appropriate.

(2) The requirements for post-observation conferences set forth in subsection (c)(10) of this section may be met through the summative conference, provided the appropriate appraisers are present and time requirements are met.

(3) Any documentation collected after the summative conference but before the end of the required days of instruction for students during one school year may be considered if it will affect the teacher's domain and overall summary performance scores. Another summative conference shall be held to inform the teacher of the changes.

§149.44 Teacher Appraisal Instrument, Scoring Procedures, and Forms.

(a) The State Board of Education shall develop and approve the Texas teacher appraisal instrument, which shall include the domains and criteria listed in §149.42 of this title (relating to Teacher Performance Criteria) and indicators for each criterion.

(b) The State Board of Education shall develop and approve scoring procedures which guarantee that each teacher, at the close of the appraisal process, receives a performance score for each domain and an overall summary performance score.

(1) Each appraiser will score Domains I-IV.

(2) The teacher's supervisor shall score Domain V. In the evaluation of Domain V the teacher is assumed to have credit for all indicators unless the teacher's supervisor has documented otherwise. In the event that the teacher's total score for the year in Domain V is less than meets expectations, the other appraiser shall review Domain V documentation and independently score Domain V for that teacher.

(3) An appraiser must document evidence on the written record when:

(A) absent or below expectation is recorded for any indicator; and

(B) exceptional quality credit is granted for any criterion.

(4) The domain credit totals issued at the close of each appraisal period by the teacher's supervisor reflect the results of the formal observation issued on written records and the additional appraisal documentation gathered during the appraisal period. The domain credits issued by the teacher's supervisor have an assigned weight of 60% of the domain total.

(5) The domain credit totals issued at the close of each appraisal period by appraisers other than the teacher's supervisor reflect the average of domain credits issued on each written record completed by the other appraisers during the appraisal period. The domain credits issued by other appraisers have an assigned weight of 40% of the domain total.

(6) For teachers qualifying for one appraisal each school year, the teacher's supervisor will determine the overall summary performance score at the end of the appraisal period. For teachers receiving two appraisals each school year, the teacher's supervisor will combine the results of the first and second appraisal to determine the overall summary performance score.

(7) Scoring of the teacher's performance is done in accordance with the Texas Education Code, §13.304, and is based on the summary domain credits issued each appraisal period by the teacher's supervisor and the other appraiser(s). The State Board of Education shall establish the standards for conversion of summary domain credits to domain performance scores of:

(A) unsatisfactory;

(B) below expectations;

(C) meets expectations (satisfactory);

(D) exceeding expectations; and

(E) clearly outstanding.

(c) The forms approved by the State Board of Education shall be used by each school district to record observation and documentation results, teacher assessment of instructional goals and outcomes information, professional growth plans, and final performance scores.

Adopted by the State Board of Education, May 1989.

APPENDIX B
Contractor Review of Published Literature

1.0 REVIEW OF CURRENT LITERATURE AND RESEARCH

1.1 Introduction

This literature review provides information for identifying successful features of effective appraisal systems and for assessing the Texas Teacher Appraisal System (TTAS) in comparison to other systems of appraisal. Discussions of appraisal systems and practices in education and in other industries and locales are part of this literature review. The organizational framework for the research reported in this literature review is based on the major issues identified for the study:

- Performance Appraisal Systems and Practices
- Content and Indicators in the Appraisal Instrument
- Scoring Procedures and Scales
- Process of Observation and Appraisal
- Qualifications of Appraisers

These five areas parallel the specific research areas targeted in the TTAS surveys and interviews.

1.2 Performance Appraisal Systems and Practices

This section summarizes recent literature and research on change in performance appraisal processes and outcomes, equity and fairness in appraisal processes, implementation requirements and problems, changes in perceptions and attitudes about appraisal, and linkage of appraisals to employment and rewards.

1.2.1 Change In Performance Appraisal Processes and Outcomes

In business and industry, the use of formal appraisal systems has been steadily increasing and frequently is combined with merit pay or pay-for-performance programs.

Surveys of pay practices indicate that more than 80 percent of U.S. companies have a merit pay program with an annual performance review and pay increase amount dependent upon a management policy decision (Hills, Scott, Markham, and Vest, 1987). In Canada, similar bonus programs were in use in 1985 by 92 percent of 306 large and medium-sized companies surveyed by the Conference Board of Canada (Cornell, 1986).

The U.S. Department of Education's School and Staffing Survey of 1988 showed that about 300,000 of the nation's 2.2 million public school teachers were receiving incentive pay in career ladder programs. When all teachers were asked whether they favored incentive programs, about 70 percent favored career ladder programs; 64 percent supported group merit bonuses; and 53 percent endorsed individual merit pay.

According to the 1990 Southern Regional Education Board (SREB) Career Ladder Clearinghouse Report, 25 states across the nation are funding teacher incentive programs that include career ladder or merit programs. Incentive programs that have been mandated over the past two or three years have generally focused on school incentives. The national status of incentive programs in 1990, as reported by SREB, is summarized in exhibit 1-1 (Cornett, 1991).

In another study commissioned by SREB, 12 SREB states reported evaluation programs for teachers (French, 1990). The purposes of teacher evaluation in the 12 states are summarized in exhibit 1-2. Eleven of the 12 states reported having evaluation systems for beginning teachers. Eleven states also reported evaluation systems for continuing (experienced) teachers. Three states (North Carolina, Tennessee, and Texas) use evaluation data to determine career ladder status or merit pay.

Both negative and positive effects emerge from pay-for-performance systems. The increase in productivity associated with merit raises supports the continued use of such

BEST COPY AVAILABLE

EXHIBIT 1-1
INCENTIVE PROGRAMS — 1990

	Local Initiative	Pilots with State Funding and/or Assistance	Full Implementation of State Program	State Program Under Development	Discussion No Legislative Action Pending	Type of Program
Alabama						
Alaska	X					Teacher Incentive
Arizona			X			Career Ladder
Arkansas						
California			X			Mentor Teacher
Colorado		X(1)		X(2) (Not funded)		(1) Teacher Incentive (2) Teacher/School Incentive
Connecticut			X(1)	X(2)		(1) Mentor Teacher (2) Career Development
Delaware						
Florida			X			School Incentive (2 Programs)
Georgia				X (Not funded)		Career Ladder
Hawaii				X		Mentor Teacher
Idaho			X(1)	X(2)		(1) Mentor Teacher (2) Career Compensation
Illinois	X					Teacher Incentive
Indiana	X(1)		X(2) X(3)			(1) Career Ladder/ Development (2) Mentor Teacher (3) School Incentive
Iowa			X			Teacher Incentive/ School Transformation
Kansas	X					Teacher Incentive
Kentucky				X		School Incentive
Louisiana		X(1)		X(2)		(1) Career Options (2) School Incentive
Maine			X			Tiered Certification
Maryland	X					Career Development Incentive
Massachusetts			X(1)	X(2)		(1) Teacher Incentive (2) Mentor Teacher
Michigan	X					Teacher Incentive
Minnesota	X(1)	X(2)				(1) Teacher Incentive (2) Mentor Teacher
Mississippi				X (Not funded)		School Incentive
Missouri			X			Career Ladder
Montana	X(1)				X(2)	(1) Teacher Incentive (2) Mentor Teacher
Nebraska						
Nevada	X					Teacher Incentive
New Hampshire	X					Teacher Incentive
New Jersey			X			Teacher Incentive
New Mexico					X	Teacher Incentive
New York		X(1) X(2)	X(3)			(1) District/School Incentive (2) Mentor Teacher (3) Teacher Incentive
North Carolina			X			Career Ladder/ Differentiated Pay
North Dakota	X				X	Career Development
Ohio		X		X		Career Ladder
Oklahoma				X		Teacher Incentive
Oregon		X				Career Development Mentor Teacher
Pennsylvania			X(1) X(2) X(3)			(1) Career Development (2) Mentor Teacher (3) School Incentive
Rhode Island					X	Mentor Teacher
South Carolina			X(1) X(2) X(3)			(1) Teacher Incentive (2) Principal Incentive (3) School Incentive
South Dakota	X			X		Mentor Teacher
Tennessee			X			Career Ladder
Texas			X(1) X(2)			(1) Career Ladder (2) School Incentive
Utah			X			Career Ladder
Vermont						
Virginia	X					Career Ladder/ Teacher Incentive
Washington		X(1)	X(2)			(1) School Incentive (2) Mentor Teacher
West Virginia				X		Mentor Teacher
Wisconsin	X(1)(2)				X(1)	(1) Mentor Teacher (2) Teacher Incentive
Wyoming	X					Career Ladder/ Teacher Incentive

Southern Regional Education Board

BEST COPY AVAILABLE

EXHIBIT 1-2

. PURPOSES OF EVALUATION SYSTEMS IN SREB STATES

	STATE											
PURPOSE	AL	AR	FL	GA	LA	MS	NC	OK	TN	TX	VA	WVA
Improvement	B C	B C	C	C	B C		B C		B C	C		B C
Certification			B	B	B	B	B	B	B C		B	
Re-employment	B C	B C					B C	B C	C	C		B C
Salary Increase						B C						
Career Ladder/Merit							C		C	C		

B = Beginning Teacher System
C = Continuing Teacher System

FRENCH/SOUTHERN REGIONAL EDUCATION BOARD, 1990

BEST COPY AVAILABLE

programs even though a true pay-for-performance system takes several years to implement properly. (Rollins, 1987; Schuster & Zingheim, 1988). To receive top performance from their employees, employers must be convinced that performance does make a difference (Grant, 1988). In addition, the different purposes of appraisal systems effect their impact (Locher, 1988), and the link between appraisals and productivity depends directly on the effectiveness of the appraisals (Buzzotta, 1989).

Much concern arises, however, over the impact of poorly designed and implemented pay-for-performance programs. The possible negative effects include (Brandt, 1990):

- lower morale and self-esteem
- unhealthy and cut-throat competition
- competition among teachers
- changes in relationships between teachers and principals
- undermining of collegiality among teachers (jealousy)
- lower teacher morale
- threat to job security

Among the possible positive effects of a teacher appraisal system with a linkage to an incentive pay program according to Brandt are the following:

- improvement of instruction
- attraction of higher quality teachers
- encouragement of teachers to improve
- strengthening the role of principals as instructional leaders

1.2.2 Equity and Fairness in Appraisal Processes

In court cases and other avenues for expressing employee concerns, recognition has been growing for the need to adapt procedural due process to performance appraisal systems (Pulhamus, 1989; Webster, 1988; Weiss, 1988). Lawsuits by current and former employers are becoming commonplace (Eyres, 1989). "Negligent evaluation" is one of the ways in which employees use evaluations to sue their employers (Webster, 1988).

To reduce exposure to liability claims and enhance a company's ability to defend itself, employers should (Eyres, 1989):

- make employees aware of performance criteria and standards in advance
- document performance problems regularly on appraisal forms and give employees a copy
- provide employees with relevant performance-based feedback
- make decisions that are compatible with the employee's evaluations and promptly evaluate non-productive employees
- train managers in appraisal techniques
- establish a guideline checklist

In the field of education, equity and fairness concerns in rating practices include (Brandt, 1990):

- evaluation bias (including halo tendencies)
- ambiguity of feedback
- inflation of scores
- ambiguity of job descriptions

In general, clearly written purposes, procedures outlined in policy statements, collective bargaining contracts, and evaluation manuals increase the likelihood that:

- performance expectations will be understood
- a uniform standard of judgement will be applied
- evaluations will be fair
- the evaluation results will be respected and used
- employees will have confidence in both the evaluations and the organization doing the evaluation

Further, clearly stated and fair purposes and procedures will minimize opportunities for successful legal challenges to performance appraisal (Stufflebeam, 1988).

1.2.3 Requirements and Problems of Implementation

Essential requirements for implementing an effective performance appraisal system include the use of procedures to ensure that the system has the following characteristics:

- The total system, including the appraisal form, the process for development, the organizational and management policies and procedures, and the purposes stated and actually accomplished are technically sound (Daley, 1989; Goddard, 1989).
- The performance appraisal has systemic guideposts with written performance goals and standards (Daley, 1989), including objectives of the work team and subsequently of the company (Lee, 1989, 1990; Mass & Moen, 1989).
- Performance appraisal is objective (Webster, 1988), is based upon a clear description of expected performance in terms of specific behaviors rather then vague attributes (Pelle & Greenhalgh, 1987), and is not used to justify preordained decisions (Weber, 1988).
- The performance to be appraised is clearly defined and communicated to employees in terms of expected accomplishments of the job (Geiss, 1987), observable actions which include realistic ways in which the job actually is performed in the workplace, the comprehensive range of behaviors associated with the job, and the differences in importance of different parts of the job (Liccione, 1988; Svatko, 1989).
- An appraisal form is developed assuring that the objectives targeted on the form match the actual objectives of the workplace (Goddard, 1989; Lee, 1990).
- The appraisal system emphasizes the quantitative (Kuehn, 1988), such as a numerical scale to reflect performance levels (Buford, 1988).
- Empirical assessment combined with expert judgement establishes the validity of the system (Barrett & Kernan, 1987; Pell & Greenhalgh, 1987).
- Appraisals are done carefully, truthfully, and on schedule (Webster, 1988; Lee, 1989; Phillips, 1987; Pulhamus, 1989), and are never changed retroactively (Webster, 1988).

- The appraisal has reliability (Marcoulides & Mills, 1988). Consistent appraisals are evident from candidate to candidate, building to building and district to district (French, 1991).

- Appraisal policies and procedures are established, communicated to employees, and strictly followed to assure that they are administered fairly (Lee and Webster, 1988; Weiss, 1988).

- The endorsement of the performance appraisal system by both management and employees is engaged (Lee, 1989-1990) by assuring that those affected by the system (stakeholders) have representation in the development and refinement of the system (Lee, 1989; Pelle & Greenhalgh, 1987; Weiss, 1988).

- Appraisal includes documentation of job behavior including critical incidents, criticisms or performance problems, and causes of positive performance (Bannister & Balkin, 1990; Christensen, 1990; Grant, 1988; Vinton, 1990).

- Corroborative review for each performance appraisal is included (Christensen, 1990; Edwards, 1989; Lee, 1990; Svatko, 1989; Webster, 1988).

- Employee performance is compared to specified standards rather than other employees (Goddard, 1989; Grant, 1988; Liccione, 1988, Lee, 1989; Weiss, 1988).

- Issue-specific feedback on performance results and suggestions to improve performance are included (Bannister, Brendan, & Balkin, 1990; Denton, 1987; Grant, 1988; Hoevemeyer, 1989; Nelson-Horchler, 1988; Pritchard, Roth, Roth, Watson, & Jones, 1989; Pulhamus, 1989).

- The appraisal system is designed to limit the cognitive demands (such as memory capacity, intelligence, orientation to detail) placed on the appraisers in order to match the requirements for accurate appraisal to the capabilities of the appraisers (Gaugler & Thornton, 1989; Heneman, Wexely & Moore, 1987). Accuracy is improved by limiting the number of dimensions appraisers are required to process (Gaugler & Thornton, 1989), limiting the number of ratings for each dimension (Lupton, 1989; Fox 1987-88), assuring familiarity with the job to be appraised (Hahn & Dipboye, 1988; Heneman, Wexely & Moore, 1987), and establishing a frame-of-reference for evaluation (Daley, 1987).

- Year-round periodic appraisals are provided as part of the total system (Buzzotta, 1989; Fox, 1987-88).

- Preparation for the appraisal is included such as pre-planning sessions and advance notification of the appraisal (Lawrie, 1989).

- Training in how to use the system is included (Hahn and Dipboye, 1988; Sims, Veres & Heninger, 1987).
- System checks are available for subsequent reevaluation and refinement of the total performance appraisal system (Hills, Scott, Markham, & Vest, 1987; Lee, 1990, Pelle & Greenhalgh, 1987; Rollins, 1987).
- Salary review is separated from performance appraisal (Cornell, 1986; Hall, Posner & Harder, 1989).

The literature includes reference to warning signals which indicate potential problems with the appraisal system. These "red flags" in a performance appraisal system include (Phillips, 1987):

- identical performance ratings for all employees
- identical standards applied to all employees in all work groups and all situations
- identical performance ratings and comments for the same employee from one appraisal period to the next
- disagreements between the appraiser and the supervisor on the date of the last appraisal
- reports of failed attempts to discharge poor employees
- evidence that promotions have been made based upon false information

Specific problems reported in the literature concerning performance appraisal systems include:

- excessive focus of management efforts on the poor performance of employees (Pulhamus, 1989)
- basing an appraisal on factors other than job performance including bias, inequity, dissonance between stated and actual practices, political rather than performance-based decisions, breach of contract, and discrimination, (Grant, 1988; Hills, Scot, Markham & Vest, 1987; Liccione, 1990)
- tolerance of poor performance (Grant, 1988)
- awkward or uncomfortable appraisal process issued (Grant, 1988)

- misunderstandings or lack of communication about expectations (Denton, 1987)

- absence of correlation between performance and pay (Cook, 1989; Geiss, 1987, Liccione, 1990; Thayer, 1987)

- merit pay limited to only a few employees (Thayer, 1987)

- unfair and unreliably administrated programs (Thayer, 1987).

The Joint Committee on Standards for Evaluating Educators has suggested the following standards and procedures for personnel evaluations (Stufflebeam, 1988):

- Implement institutional policies that require all employees be subject to systematic evaluation.

- Develop clearly written guidelines for implementing personnel evaluation policies.

- Ensure that the guidelines address all elements for acceptable evaluations.

- Concentrate the guidelines on important job-related issues, and avoid listing rules for trivial aspects of the job or matters unrelated to the requirements for successful job performance.

- Make the procedures sufficiently specific to guarantee shared understanding of the purposes, procedures, and substance of evaluation.

- Identify the performance reasonably expected of an employee in order to be competent and successful on the job.

- Require that appropriate weights be assigned to each evaluation criterion explicitly and in advance of the evaluation.

- Involve the board and staff in the development and periodic review of the policies and guidelines.

- Ensure that the guidelines meet all local, state and federal legal requirements concerning employment decisions, such as the state's teacher certification law or a city's non-discrimination ordinance.

- Explain the plan of personnel evaluation to all employees at least annually and when changes in the evaluation are to be implemented.

- Assure the consistent enforcement of the written evaluation standards.

- Provide a plan of progressive discipline, such as an oral warning, a written warning, disciplinary layoff, and discharge.
- Define types of evaluation findings likely to lead to termination.
- Apply the standards with a concern for the human dignity and worth of the persons involved.
- Establish viable review or reevaluation, problem-solving, and appeal procedures to protect all involved in the evaluation.
- Establish a process for periodic review and revision of the evaluation procedures and guidelines.

The critical features of appraisal systems which are used to assess the quality of teaching performance include (Brandt, 1990):

- The people selected to design and operate the program must be among the best available.
- The school administration and school board must assign high priority to the planning and monitoring of the program in order to accomplish long-term objectives.
- The administration and board must clearly choose and support with policy and training the long-term objectives concerning the roles of principals and teachers, and the structure of the staffing.
- Policy-makers must engage the participation and support of students, parents, citizens, and professional associations, with sufficient sensitivity to the needs of the different stakeholders to allow them time to share in the responsibility for the program and for making program improvements.
- The program for teacher performance assessment for incentive pay should include features from more than one model.
- Voluntary, rather than mandatory, participation in a teacher appraisal system with a linkage to incentive pay should be allowed.
- Multiple criteria, multiple judges, and multiple data sources for teacher appraisal should be used.
- Variables and performance indicators must be clearly described, illustrated, and widely distributed.
- Data collection must be treated as a process separate from evaluation.

- Evaluators should be trained in the collection and judging of data. Further, the competency of evaluators in their use and knowledge of the evaluation system must be demonstrated.

- Research and measurement specialists, outside consultants, and third party evaluators should be used to conduct continuous monitoring, research, evaluation, and to provide technical assistance.

- Annual reports of participation patterns and surveys of teachers, administrators, and others should be prepared. The results should be used to modify and improve the system.

- Sources for long-term funding of a teacher appraisal system with linkage to an incentive program should be identified and procedures to keep participation rates congruent with multi-year cost projections must be developed.

1.2.4 Changes in Perceptions and Attitudes

In the team- and service-oriented corporations of the 1990s, the opinions of employees and customers about management, including appraisal systems, are of increasing importance (Nelson-Horchler, 1988).

The attitudes of employees toward performance appraisal systems, whether or not linked to merit pay, depend directly upon the presence of the following processes and events:

- the adequacy of and the way in which information is communicated concerning the appraisal system, performance results, and pay if linked to performance (Bannister & Balkin, 1990; Cornell, 1986; Denton, 1987; Hoevemeyer, 1989; Korinek & Thobe, 1989; Lee, 1989)

- the methodology of appraisal including consistency of practice with prescribed procedures, objectivity of appraisal, adequacy of preparation for appraisal, fairness and reliability of administration, comfort of appraisal process (Buzzotta, 1988; Cornell, 1986; Denton, 1987; Korinek & Thobe, 1989; Lee, 1989; Shelley, 1987; Thayer, 1987)

- the form and amount of recognition and reward for performance and the relationship of appraisal results and rewards to performance (Cook, 1989; Geiss, 1987; Grant, 1988; Shelley, 1987; Thayer, 1989)

- the standards of performance with a preference for avoiding comparison with other employees (Grant, 1988) and for having clearly written goals and standards (Geiss, 1987; Lee, 1989)

- quality of feedback from appraisals including an explanation of the causes of both positive and negative results, and the separation of discussion of appraisal results from salary issues (Bannister and Balkin, 1990; Cornell, 1986)
- the congruence of evaluation criteria and procedures with new or proven practices (French, 1990)
- the presence or absence of monitoring for the system (Lee, 1989).

In the field of education, teachers and others involved in teacher appraisal have expressed a number of concerns about teacher appraisal and about appraisal linked to incentive programs. These concerns include (Brandt, 1990):

- whether incentive pay is worth the extra effort, time, stress, or paperwork
- whether most teachers are effective despite obvious differences in teaching style and personality and whether great teaching can be distinguished from good teaching
- what criteria to consider and a precise description of the criteria for successful teaching
- solid documentation of in and out-of-class performance areas
- a precise description and explanation of any perceived shortcomings in teaching practices
- assistance in achieving top ranking
- an explicit explanation of reasons for a teacher's failure to receive top ranking
- measurement of the impact on students
- fair and equal consideration given vast differences in teaching assignments, grade levels and subjects
- how much to weigh a principal's judgement versus an assessment by other individuals
- capability of evaluators
- kind and quality of evaluator training
- minimal publicity for those receiving awards or those who are promoted

- avoidance of teacher-to-teacher comparisons
- use of low inference coding versus high inference coding
- partiality of evaluators, especially principals
- the credibility of recipients of incentive awards.

Other concerns expressed by French (1990) are:

- the restriction of data collected to a limited range of practices or models of teaching
- the relationship of teacher practices to student outcomes.

1.2.5 Linkage to Employment and Rewards

Incentive or performance-for-pay systems often have led to improvements in performance of employers as significant as "10-20 percent or greater" (Pritchard, Roth, Roth, Watson & Jones, 1989). One study showed that the merit pay system in a bank resulted in profits improving by 10 percent while return on assets soared to greater than 100 percent (Schuster & Zingheim, 1988). Rollins (1987), in a review of the pros and cons of merit pay, also concluded that merit raises can increase productivity. Some data indicate, however, that raises below four percent produce feelings of unfairness (Korinek and Thobe, 1989).

To avoid legal problems, appraisals should be specific and provide support for salary or bonus decisions. But, as opportunities to move up decrease, development appraisals need to be separated from salary increases and promotions (Nelson-Horchler, 1988). Most of the 36 high technology firms surveyed in California's Silicon Valley, tended to separate the salary review discussion from the performance appraisal discussion (Hall, Posner & Harder, 1989; Cornell, 1986).

In one survey of the perceptions of workers as to the linkage between performance and pay, "only 28 percent [of those evaluated] have a positive view, and only one-third receive regular performance feedback", (Anonymous, 1988). Overall, the causes for dissatisfaction with such systems include the same causes of dissatisfaction associated with negative attitudes of performance appraisals, such as lack of a direct relationship of pay to performance, unfairness or inequity in consistency of process and results, organizational problems, bias and politics.

Incentive plans that work, however, as reported in the banking system, include the following characteristics (Schuster & Zingheim, 1988):

- incentives are tied to performance
- credit quality is as important as quantity as a measure of performance
- incentive compensation is reserved for front-line officers rather than executives
- incentive compensation is a larger proportion of the total cash compensation.

In the field of teaching, appraisal practices are also frequently linked to merit pay. In its third year, the Tennessee Career Ladder program was showing positive linkages of employment and rewards. Furtwengler (1987) reported that exciting things are happening as a result of the Career Ladder Program:

- teachers are receiving higher salaries and incentives for outstanding performance
- Career Ladder teachers are serving as mentors to beginning teachers
- the majority of Career Level II and III teachers are electing to work extended contracts and to provide new learning opportunities for students
- staff development opportunities for teachers are increasing and are being geared to meet developmental needs
- student achievement scores in Tennessee are improving.

While comprehensive evaluations of teacher appraisal and incentive programs show positive results -- including improvements in student achievement and changes in how districts involve teachers in improving curriculum and instruction -- they also reveal problems. In many school districts, the existing organizational structure may prevent real change from occurring. Incentive programs that focus on individual performance or elevate the status of one teacher over another continue to meet with resistance. Linking rewards for individual teachers to student achievement is problematic and has received little attention except in a few school districts. Programs that provide extra pay for extra work or focus on school-wide rather than individual incentives are increasingly popular. These programs are less costly, can be put into place more quickly, and do not require fundamental changes in how teachers and principals go about their work (Cornett, 1991). On the other hand, the smaller awards to all teachers, or to the school, may provide less overall incentive than individual award systems (Brandt, 1991).

1.2.6 Major Features of Selected Teacher Appraisal Systems

A number of state and district programs are designed to recognize and reward individuals who consistently perform in an outstanding manner with promotion to a higher rank on a career ladder and a substantial salary supplement.

Differences among these programs include (Brandt, 1990):

- the type of awards (status vs. salary)
- the size of awards (hundreds vs. thousands of dollars)
- criteria, measures, and standards (what teachers must do and how and by whom they are judged)
- primary purpose (to reward only superior teachers vs. to encourage all to improve vs. to increase student achievement vs. other goals).

Other factors for promotion on a career ladder include extra responsibility; extra hours; teacher attendance; written skills; student, principal, and peer ratings or responses to questionnaires; and student learning. Career Ladder programs in Tennessee, North Carolina and Texas have been reported in the SREB research (Cornett, 1991). Exhibit 1-3 highlights the major features of these three state programs.

EXHIBIT 1–3

MAJOR FEATURES OF SELECTED STATE CAREER LADDER TEACHER EVALUATION SYSTEMS

	Tennessee ***Career Ladder Levels II & III***	***North Carolina*** ***1989 Pilot Districts***	***Texas*** ***(TTAS)***
Data Sources	Teacher candidate Students Principal Peer evaluator (2–3)	Teacher candidate Principal Assistant Principal Peer evaluator	Teacher candidate Observed class reactions Varies by district, primarily administrators
Instruments	Professional Development and Leadership Summary Class observations Dialogues (Interviews) Professional knowledge test Reading/writing tests Student questionnaires Principal questionnaires Peer evaluator team consensus	Class observations Local forms for out-of-class activities in some districts	Class observations
Class observations	6 in a scheduled week	4 for career ladder promotion or 3 to maintain status, only 1 must be scheduled	Scheduled 4 CL I, 2 CL II, III*
Preobservation Conference	Required but not immediate feedback	Required for scheduled observations	Optional, varies by district
Postobservation Conference	Required	Required	Required of supervisor -- not of other appraiser, unless unsatisfactory
Objective Description	Required procedure	Required procedure	Encouraged, no statewide prescribed procedure
Evaluation Record	Required format (scores plus printout of strengths and weaknesses)	Required format	Required format
Out-of-district observors/evaluators	Required (2–3)	Not used	Not used
Basis for Professional Improvement Planning	Limited by delayed feedback, but feedback is specific	Feedback on specific skills	Plan required only for unsatisfactory performance or from teacher request

EXHIBIT 1-3

MAJOR FEATURES OF SELECTED STATE CAREER LADDER TEACHER EVALUATION SYSTEMS

	Tennessee Career Ladder Levels II & III	*North Carolina 1989 Pilot Districts*	*Texas (TTAS)*
Evaluator Training Requirements	150 hours, pass competency test on all instruments	54 hours	90 hours, test on TTAS criteria/indicators
Nature of Criteria/Teaching Skills	Effective teaching research base includes five observed areas of behavior	Effective teaching research base	Effective teaching research base
Involvement of Peers	State-funded cadre of teachers to serve as outside evaluators to local districts	State-funded peer observer positions filled by local-employed teachers	Limited to lead teachers or department chairs
Student Outcome Data	New requirement, not fully developed	Not included	Observed student reaction only
Responsibility for Final Evaluation	Team of 3 outside evaluators (or 2 plus principal if requested)	Principal/Asst. Principal (increasing shared with peer evaluators)	Supervisory administrator
Score Distribution by Career Ladder Levels	Cone, wide distribution range with clear differences between levels	Truncated cone, two thirds of qualified teachers at highest implemented level (CL II)	Diamond, varies by district, ceiling effects
Special Features	Multi-source, multi-instrument data, outside evaluator system for CL II & III	Heavy involvement of local, state-funded teacher evaluators	Required documentation of teacher behaviors for extra quality points, use of student reactions in observation system

* CL = career ladder

The major differences between the Texas Teacher Appraisal System (TTAS) and the Tennessee and North Carolina Career Ladder teacher appraisal systems include:

- Class or student reactions are observed in Texas and Tennessee.
- Observers are teachers in Tennessee and North Carolina. In North Carolina full-time teacher/observers are selected to serve two or more years. Principals and other administrators are observers in Texas.
- Texas has one observation instrument, North Carolina permits local forms for out-of-class activities, and Tennessee uses multiple data sources (observations, interview, etc.).
- All Texas classroom observations are scheduled, only one must be scheduled in North Carolina and in Tennessee both scheduled and unscheduled observations are conducted.
- Pre-observation conferences are optional in Texas, while they are required in the other states.
- Texas does not require a specific kind of observation record.
- The Tennessee Career Ladder requires observers from outside of the district, while the other two states do not.
- Specific feedback is required in Tennessee and North Carolina but is required only for unsatisfactory performance in Texas.
- Evaluation training varies from 84 hours in North Carolina to 150 hours in Tennessee.
- Tennessee and North Carolina provide state-funded peer teacher observers.
- Tennessee has a cone shaped core distribution of scores with distinctive performance levels.
- Texas has a provision for extra quality points to be awarded by the evaluator.

Orange County, Virginia has another program with special features which over the years receives support from teachers with over 98 percent participation and 90 percent receiving small bonuses. Organized around "twelve professional areas and fifteen procedures for effective teaching" are eight to twelve generic performance indicators and examples. Special features are as follows (Brandt, 1990):

- Participating teachers each choose one broad area to concentrate on for a nine-week period with up to three areas a year.
- During that time trained peer observers make two classroom visits on days selected by the teacher. Observers record in objective, narrative language teacher behaviors and actions that exemplify the performance indicators in the target area. Pre-conferences precede the observations for discussion of performance indicators, and post-conferences allow the teacher a chance to review the observation report for completeness and accuracy before it is sent to the principal for rating. Thus, peer observers provide the documentation, and the principal judges what the record indicates in relation to the standards. Reviewing the two or more observation reports, the principal rates the teacher's performance with each targeted practice as "insufficient," "competent" or "proficient".
- A rating review by the central office can be requested if a teacher is dissatisfied with a rating.
- The program maximizes the positive and, with very few exceptions, ignores the negative. Teachers can select the areas in which to be observed and judged. Observers only record what is right and never what is wrong.
- The program selects, trains, and uses a cadre of highly respected teachers to serve as peer observers for providing documentation and formative assistance for other teachers. Approximately 15 percent of all the teachers serve this function in any given year on a released time basis while continuing to teach their own classes as well. They receive 30 hours of training and from $500 to $1,600 extra pay.
- The program restricts the role of peer evaluators to observer-recorder which provides the evaluator with better documentation and provides the evaluator time to work with teachers who need special assistance.
- The program receives favorable reactions from teachers with 70 to 80 percent indicating the observation reports make them feel better about the job they are doing and more than half of them saying they would recommend the pay program to other teachers. Importantly, half of them typically report that they did some different things the following year to become proficient in at least one of their targeted practices.

The Danville, Virginia program had a number of special features that distinguished it from other programs (Brandt, 1990).

- There is no initial rung in the career ladder, as with the Tennessee plan, where the vast majority of the teachers will participate in a salary supplement. The rungs above basic salary levels on the Danville ladder are clearly for those who are above average in overall performance.
- The quality of teaching performance is the primary basis for promotion and extra pay, not qualification for extra responsibilities nor willingness to work longer hours or extra days.
- Performance assessment data are derived from a variety of sources: observations by administrators and teacher evaluators, conferences with career ladder selection committees, student performance data collected early and late in the year, parent and student surveys, and teacher records.

Other special features include:

- teacher selection of one of the three members of the review team
- clear specification of 45 performance criteria and instructional variables
- special training for all evaluators
- a coordinating committee of teachers and administrators to enhance procedural and judgmental consistency across schools, grade levels and classes, review teams, and candidates
- evidence that significant student learning occurs
- inclusion of librarians and speech therapists

1.3 Content and Indicators in the Appraisal Instrument

1.3.1 Adequacy and Appropriateness of Measures

Most white-collar workers have no history of performance measurement, and therefore they associate measurement with management practices aimed at lower paid blue-collar workers. After eight years of experimentation, Boyett and Conn (1988) found that white-collar measures can be developed and when used in the proper way, significant performance gains are possible. Important findings included:

- Results rather than activities of white-collar workers must be measured.
- White-collar work should be classified into broad categories such as processing and support functions or quick response, recurring and long-term projects.
- Group or team measures are used, rather than individual measures, because white-collar workers usually work on group projects.
- Consensus building is utilized to involve white-collar workers in the design and development of measures.
- A family of indicators such as quality, timeliness and cost for each measure is used.

An appraisal system must be built upon clear job-related performance and promotion standards which are applied consistently and are documented (Metz, 1988).

- The performance to be appraised must be clearly defined in terms of expected accomplishments on the job, observable actions, realistic ways in which the job actually is performed in the workplace, the comprehensive range of behaviors associated with the job, and the differences in importance of different parts of the job.
- The objectives targeted in the instrument must have validity and match the actual objectives of the workplace.
- Empirical assessment combined with expert judgement must have been used to establish validity.

The major sources used for defining domains and criteria in teacher evaluation programs include (French, Holdzkom and Kuligowski, 1990):

- research literature on effective teaching
- job descriptions
- job analysis
- teacher consensus
- state policies/legislation

The primary areas included in teacher appraisal programs include (Ibid.):

- instructional planning
- instructional delivery
- classroom management
- student assessment
- basic communication skills
- personal characteristics
- professional responsibilities
- professional leadership.

French, Holdzkom, & Kuligowski report on the individual criteria for each of these primary areas of appraisal in beginning and continuing teacher education programs across 12 states.

1.3.2 Assessment for Different Levels of Employees

The research in business and industry indicates the need for separate appraisal instruments and processes for different levels of employees when they have different responsibilities by virtue of the requirement for matching job requirements.

In the field of teaching, research-based instruments used to observe, record, and evaluate behaviors associated with effective teaching are "desirable for novices and advanced beginners to learn and use but great numbers of effective teachers exist who do not do those things at all or who do them in some different form than that prescribed by researchers" (Berliner, 1988). After examining the findings from a number of studies Berliner reached the following conclusion:

> *Instrumentation that is perfectly sensible to use with novices may not be suitable for proficient and expert teachers who have found their own ways of accomplishing the tasks of teaching. These teachers are more intuitive and holistic in their ways of teaching, and their classroom behavior may look different than that of novices or advanced beginners.*

French, Holdzkom & Kuligowski (1990) report that few states, including Texas, have taken this conclusion into account.

Some other relevant findings in the literature include the linear relationship between test scores and job performance when validity tests are used for the appraisal (Giffin, 1989) and the predictive validity of task-performance measures, which actually are criterion-referenced tests usually administered in a real-world or simulated context (Campbell & Armstrong, 1988).

Whether using valid tests or job-reference task performance measures or other forms of appraisal instruments for different levels of employees (i.e., for master teachers), the same requirements for validity and legal defensibility must be met. The performance to be appraised for each teaching level must be clearly and realistically defined in terms of expected accomplishments and observable actions across the full range of job requirements with objectives which clearly match the actual objectives of the workplace. In addition, validity must be established for each level through empirical assessment combined with expert judgement.

1.3.3 Criteria and Standards for Appraisal

In the earlier section on the requirements and problems of implementation (Section 2.2.3), the following key features were identified as being essential components of the content and standards in the appraisal instrument:

- systematic guideposts with written performance goals and standards
- objectively written, clear descriptions of expected performance in terms of specific behaviors
- descriptions of observable actions which reflect realistic ways in which the job actually is performed in the workplace, the comprehensive range of behaviors associated with the job, and the differences in importance of different parts of the job

- targeted objectives which match the actual objectives of the workplace

- quantitative values to reflect performance levels

- a method and format for documentation of job behavior, including critical incidents, criticism, performance problems, and causes of positive performance

- a method and format for the appraisal which limits the cognitive demands placed on appraisers in order to match the requirements of the appraisal instrument and process to the capabilities of the appraisers. Appraisers should be able to reliability distinguish between categories and to explain the results.

Behavioral anchors, which contain examples of "good" and "bad" behaviors to focus the attention of the appraisers are another approach for clarifying criteria or standards of performance (Attoy & McIntrye, 1987; Fox, 1988/89; Tziner, 1989).

The criteria for appraising teacher performance in 12 southern states, as reported by French, Holdzkom, & Kuligowski (1990), demonstrate agreement in definitions of broad performance areas. However, the agreement lessens as more finite indicators of performance are defined and only nine programs in five states attempt to directly relate student outcomes to teacher evaluation. In their study of teacher evaluation in SREB states, they conclude:

- Criteria used to judge teacher performance in the twelve states participating in the study are very similar at one level. Most of the teacher evaluation programs submitted for comparison assess teacher planning, delivery of instruction, assessment of student progress, classroom management, student involvement, teacher basic communication skills, classroom climate and teacher interpersonal skills. However, there is substantially less agreement about the definitions and clustering of performance indicators or teacher behaviors that constitute these areas or domains of competence.

- Where evaluation programs for both beginning and experienced teachers exist within the same state, they tend to focus on the same assessment criteria. However, data pertinent to the criteria may be gathered in different ways. (For example, one observation may be used in evaluation of a continuing teacher, but three are conducted in the state's beginning teacher evaluation. Or, an interview process may be used to collect information about planning from continuing teachers while beginning teachers submit lesson plans for review).

- In deriving teacher evaluation criteria, there has been relatively little attention given in SREB states to the emerging knowledge base about differences between beginning and experienced teachers. Few evaluation programs use different criteria or weight criteria differently for the two groups.

- Only five evaluation programs in four states report assessment of teacher innovative practices (e.g., cooperative learning). State representatives to the SREB study report that their evaluation programs try not to inhibit innovative practices, but that these programs do not reward such practices.

- Only nine programs in five states attempt to directly relate student outcomes to teacher evaluation.

- Teacher practices demonstrated to have relationship to school effectiveness (sharing ideas/materials, initiating activities/projects, assisting peers) are included in teacher evaluation in only five states.

- A concern for continuous teacher professional growth and instructional improvement is exhibited in the inclusion of professional development assessment criteria in thirteen evaluation programs developed in nine of the twelve states participating in the study.

1.3.4 Appraisal of Outcomes

In business and industry, the standards for appraising performance have included outcomes measures such as (Brandt, 1990):

- piecework pay, sales commissions, executive bonuses for production quotas

- evaluation by supervisors, senior officers, test results, completion of special training, time in rank, time in service

In the field of education, many parents and community leaders feel teachers should be judged by how much their students learn. However, it is difficult to isolate the influence teachers have on students in consistent, reliable ways (Brandt, 1990).

Few comprehensive assessments of performance appraisal systems that appraise outcomes have been made. However, the state of Arizona has extensive information on its program. In establishing the pilot career ladder program, the Arizona Legislature mandated an extensive evaluation of its effect over the life of the pilot projects. The research results led the Legislature to expand the project in the state. The five-year research and evaluation project found that despite the "diversity of participating districts and researchers," the program has had a positive effect on student achievement in the districts participating in the program. The study reported that (Cornett, 1991):

- Students taught by career ladder teachers are making significant gains on achievement (outpacing those taught by non-career ladder teachers).

- Student achievement depends more on teacher performance than years of experience in the classroom.

- Districts in the Career Ladder Project report that, due to their participation in the project, they have:

  - better alignment of curriculum

  - teachers setting higher level learning objectives

  - administrators who are more involved in day-to-day classroom teaching

  - better communication and sharing among teachers

  - more locally developed learning methods and materials

  - greater opportunities for teachers to participate in in-service and peer coaching (teachers not in the program still randomly accumulate college and in-service credits)

  - an increased number of "teacher leaders" in the districts.

An extensive third party validation study was recently completed on the Tennessee career ladder evaluation system (Baker, et.al., 1990). This study was mandated by legislation to insure that educators in the program were receiving fair and objective assessments of their competency. The study found (Reported in Cornett, 1991):

- Sound procedures were used to identify the domain and competencies that are assessed.

- The competencies and indicators are based on effective teaching results and consensus by Tennessee teachers.

- Procedures for selecting evaluators are sound and training is adequate.

- A higher proportion of black teachers were employed at Career Level 1 than would be expected, and a lower proportion at Levels 2 and 3. Female teachers reached Level 3 at a greater success rate than did male teachers. The percentage of teachers reaching the higher levels did not vary by region of the state, but did vary on per capita income of the county (the higher per capita income, the greater the percentage of teachers on upper levels.) The report indicated that further study was necessary to develop reasons for these findings.

- "The staff has done a stellar job of conceptualizing, developing, and implementing a model teacher evaluation system that allows for valid inferences for these teaching qualities."

1.4 Scoring Procedures and Scales

1.4.1 Method for Rating

The most common measurement methods in business and industry are scales based on trait-ratings, behavior, effectiveness, and hybrids of each of these which assess both performance and potential (Goodstein, 1989). Some measures make the rating of quality, defined in a measurable way, as important as quantity as a measure of performance (Campbell & Armstrong, 1988; Schuster & Zingheim, 1988).

A scale for rating performance, such as 1-5 (Buford, Bourkhalter, Jacobs, and Grover, 1988) or outstanding, satisfactory, and unsatisfactory, (Lupton, 1989), frequently applies to written definitions or critical incidents for each performance level. Behavioral anchors also frequently accompany scales to more clearly define criteria (Gomez, Mejia, 1988; Murphy & Constans, 1987; Tziner & Latham, 1989).

Limiting the number of dimensions to the amount that appraisers can reliably distinguish and explain will support the accuracy of ratings. Although the accuracy of observations is not affected, the accuracy of performance ratings decreases when appraisers are required to observe, classify, and evaluate a large number of dimensions, (Gaugler & Thornton, 1989). French (1990) reports the most common performance measurement method for evaluating teachers is classroom observation. He also reports:

- In the SREB states participating in this study, there is heavy reliance on classroom observation as a source of evaluation data. In three states it is the only source of information used.
- Instruments and/or data collection procedures most often used in addition to observation are candidate interviews, review of administrator records and candidate self-reports.
- There is a trend in the states participating in this study toward the use of multiple data sources in teacher assessment (as indicated by the fact that nine of twelve states report the use of more than one source of information).

1.4.2 Influence of Factors Other Than Performance

Various factors may distort the accuracy of a performance appraisal, some of which have been cited in earlier sections of this report. Such factors include the quality and integrity of the appraisal form, the process for development of the appraisal system and for establishing validity of the appraisal, the organizational and management policy and procedures, and the fairness and consistency of administration of the published appraisal policies.

Even when all of the basic requirements for establishing a fair and valid appraisal form and policies are met, the accuracy of performance ratings may be influenced by additional factors. For example, Marcoulides and Mills (1988) reported the results of a study on appraisal of secretaries in which supervisor ratings were a substantial source of error variation. Nonetheless, with a rating form of ten items that contribute to the assessment of overall job performance, some supervisors still rated people using different criteria. Marcoulides and Mills also reported that often employee performance evaluation is made difficult by the evaluation instruments themselves.

Geber (1988) acknowledges that it is not possible, however, to design a performance appraisal that ensures perfect objectivity and that no matter how much an agency may spend on creating a system and providing appropriate training, managers and supervisors still tend to falsify performance ratings by deflating or inflating an employee's ratings. Further, managers and supervisors tended to place more importance on the use of appraisals as a way to motivate employees rather than on rating accuracy.

The reasons cited for deflating appraisals include:

- to encourage an employee to quit
- to create a record to justify a planned finding

The reasons included for inflating appraisals are:

- to help a worker whose performance is suffering because of temporary personal problems
- to avoid creating a permanent record of poor performance
- to avoid confrontation with hard-to-manage employees

The following recommendations are among the alternatives reported as ways to improve the rating accuracy in appraisal systems:

- Document critical incidents and job behavior at the time of the behavior or shortly after the behavior occurs (Fox, 1987/88; Vinton, 1990).
- Reduce the number of categories of performance to those that supervisors can reliability distinguish and clearly explain (Lupton, 1989).
- Use multiple sources and multiple raters (Edwards, 1989; Fox, 1987/88).
- Give credit for good form, documented in specific critical incidents, rather than placing an undue emphasis on outcomes (Fox, 1987/88).
- Rate only extreme performance, excluding the middle (Fox 1987/88).
- Define "good form" by citing specific positive and negative critical incidents (Fox, 1987/88).
- Use a coordination panel to process the evaluations (Fox, 1987/88).

1.5 The Processes of Observation and Appraisal

1.5.1 Preparing for and Conducting Appraisals

Appraisals which fail tend to be those with a lack of preparation, confusion, and overreliance on forms (Buzzotta, 1988). Employees, as well as evaluators, require preparation for appraisals, including pre-planning sessions and advance notification of the appraisal (Lawrie, 1989). Making employees aware of performance standards and criteria in advance is an important feature of a legally defensible system (Eyres, 1989).

With an effective appraisal system, the processes of preparation and follow-up can make a substantial difference in the impact on employees by assuring them that their performance actually does make a difference. Relevant strategies used in personnel management as part of follow-up to an appraisal include (Grant, 1988):

- rewarding or "talking up" performance
- coaching and counseling on performance
- posting performance results
- delegating and following-up clearly on responsibilities
- rewarding suggestions that lead to improved performance
- conducting meetings that focus on performance
- making performance satisfying

When linked to merit pay, performance appraisal must be transferred into performance management in which expected job accomplishments are clearly defined and communicated. Such performance management begins with the definition of the most important job duties and outcomes coupled with on-going daily management of people and problem solving so that results are integrated with behavior (Geiss, 1987). During appraisal interviews, positive, open, and trusting experiences can be built upon the following approaches (Denton, 1987):

- planning objectives

- allowing the employee to talk freely
- giving praise
- expressing empathy, firmness, and fairness
- determining causes and solutions to performance problems (cf. Lawrie, 1989)

A five-phase performance cycle that will enhance the performance-orientation of the system will include (Hoevemeyer, 1989):

- performance contracting which defines the priority of each job area of accountability and specific measures of performance
- information sharing with open communication
- performance recognition which rewards or compensates employee for their achievements
- annual performance evaluations which review actual achievements
- performance reinforcement including periodic evaluation with feedback

Other approaches to preparation and follow-up include ways to enhance the legal defensibility of a system such as the following (Eyres, 1989):

- Make employees aware of performance criteria and standards in advance.
- Give employees a copy of documented performance problems.
- Provide employees with relevant performance-based feedback.
- Make decisions that are compatible with the employee's evaluations.
- Promptly evaluate non-productive employees.
- Provide an avenue for the employee to comment on or dispute the performance appraisal.
- Train appraisers in appraisal techniques.
- Establish a guideline checklist.

1.5.2 Observation Time Periods and Frequency

Performance appraisal problems may result from the process used in appraising performance even when indicators are valid and clear (Marcoulides & Mills, 1988). The major findings from business and industry concerning observation time periods and frequency as part of the total process include (Kamouri and Balzer, 1990):

- Those who equally sample all persons evaluated make more accurate relative estimates than those who unequally sample employee performance.
- Those who observe a greater amount of information provide more accurate ratings than those who observe less information.

In the SREB study of teacher evaluation, French (1990) found observation periods range from 20 minutes to a full class period. The number of observation for different purposes ranged from one for tenured teachers to nine for beginning teachers. Exhibit 1-4 summarizes information on observation procedures for the 12 states in the SREB study.

1.5.3 Consistency and Monitoring of Appraisal Procedures

Consistency of procedures used by appraisers influences the accuracy and fairness of performance ratings (Shelley, 1987). A common risk in treating employees unequally is the failure to carefully follow published policies so that standards are uniform (Barrett & Kernan, 1987), consistently applied, and documented (Metz, 1988). Such consistency becomes especially important when an appraisal system is linked to merit pay because merit pay systems succeed only when employees perceive them as being fair and reliably administered. Anything less leads to a significant decline in morale (Thayer, 1987). Although the benefits of an effective appraisal system include improved overall morale as well as a dynamic and progressive work environment, an unmonitored system can lead to failure (Lee, 1989).

EXHIBIT 1-4
OBSERVATION PROCEDURES

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Alabama	B LEA option C LEA option	observers	B 6 C 4	B Class or 1 hr. C Class or 1 hr.	B Yes C Yes	Yes Yes
Arkansas	B Script C Script	observers	B 8 C 2	B 30 min. C 30 min.	B Yes (2) C Yes (2)	Yes Yes
Florida	B Sign C LEA option	observers	B 5 C LEA option	B lesson	B Yes C No	Yes Yes
Georgia	B Mod. Script C Mod. Script	observers	B 3 C 3	B 30 min. C 20 min.	B No C No	No No
Louisiana	B Mod. Script C Mod. Script	observers	B 6 C 6	B 30 min. (Elem.) Class (Sec.) C 30 min. (Elem.) Class (Sec.)	B Yes C Yes	Yes Yes
Mississippi	B Sign C Sign	observers	B 6 C 1	B Class or 1 hr. C Class or 1 hr.	B No C No	Yes Yes
North Carolina	B Mod. Script C Mod. Script	observers	B 6 C 3 C 4 (Merit)	B 20-60 min C Class	B Yes (1) C Yes (1)	Yes Yes
Oklahoma	B Script C LEA option	observers	B 9 (1st year) C 1 (tenured) B 2 (non-tenured)	B LEA option C LEA option	B No C No	Yes Yes
Tennessee	B Mod. Script C Sign	observers	B 3 (1st year) C 2 (Professional) C 6 (Career Ladder) B 2 (non-tenured)	B Class or 50 min.	B Yes C Yes	Yes Yes
Texas	C Mod. Script	observers	C 4 (candidate) C 2 (maintenance)	C 45 min.	(C) Yes/No	(C) If defic.
Virginia[1]	B Sign	observers	B 3 (minimum) B 9 (maximum)	B 35 min.	(B) No	(B) No
West Virginia[2]	B LEA option C LEA option	observers	B 4 C 2	B LEA option C LEA option	B No C No	Optional Optional

Key: B=Beginning Teacher Evaluation, C=Continuing Teacher Evaluation

[1] Texas has no state evaluation program for beginning teachers.

[2] Virginia has no state evaluation program for continuing teachers.

FRENCH/SOUTHERN REGIONAL EDUCATION BOARD

BEST COPY AVAILABLE

One example of the problems of consistency of appraisal procedures is the tendency to evaluate traits or knowledge, skills, and abilities rather than actual job performance -- a tendency which may be offset by effective training (Sims, Veres & Heninger, 1987). Inconsistent procedures also may create problems of discrimination, defamation, or wrongful discharge (Webster, 1988).

To promote consistency of appraiser procedures, monitoring guidelines should be in place to assure the following (Webster, 1988):

- Evaluations are accomplished carefully, thoroughly, truthfully, and on schedule (cf. Lee, 1989).
- Established evaluation procedures are strictly adhered to.
- Evaluations are objective with systematic guideposts (cf. Lee, 1989).
- Criticisms are documented.
- Evaluations are reviewed by at least one other person (cf. Lee, 1989).
- Evaluations are never changed retroactively.
- Evaluations are not used to justify preordained decisions.

1.6 Appraiser Qualifications and Training

1.6.1 Appraiser Background

The background of appraisers influences the accuracy of performance ratings. For example, those who give more accurate performance ratings include:

- experts compared to non-experts (Smither and Reilly, 1989; Smither, Barry, and Reilly, 1989)
- those with more information on the jobs they are rating compared to those with less information (Hahn & Dipboye, 1988; Heneman, Wexely, & Moore, 1987; Kamouri & Balzer, 1990; Smither & Reilly, 1989)
- those with training in job evaluation techniques compared to those without such training (Hahn and Dipboye, 1988)
- those who have high memory capacity compared to those with limited capacity (Heneman, Wexely, and Moore, 1987)

- those who are intelligent and detail-oriented (Heneman, Wexely and Moore, 1987)
- those who believe in the variability of human nature (Heneman, Wexely Moore, 1987)
- those with a behavioral frame of reference for ratings (Daley, 1987)
- those with enhanced opportunities to observe performance (Smither, Barry and Reilly, 1989).

Seven states (see exhibit 1-5) report that teachers are utilized as classroom observers (French, 1990). All seven of these states are included among the eight states (10 programs) requiring observer teams. Further, in all these seven states, teachers who serve as observers also serve as evaluators, a relationship that is not always present. The fact that in more than half the states participating in this study teachers serve as observers and evaluators of other teachers marks a dramatic change in teacher evaluation in the last ten years, and this procedure underscores the concern for teacher input into implementation of evaluation programs as well as their development.

Ten of the 22 evaluation systems analyzed require a mix of observers. Groups from which the observer team may be drawn include principals, assistant principals, teachers, supervisors, specialists, retired educators and higher education faculty. Nineteen programs in the ten states depend upon evaluators from the same local education agency as the teachers being evaluated. Most often (18 programs), the local superintendent or the evaluatee's building principal select the evaluator. Only Oklahoma's beginning teacher assessment program allows the evaluatees voice in evaluator selection. Persons who serve as classroom observers also serve as evaluators in 16 programs (10 states). One state (West Virginia) leaves that decision to the local education agency.

The expertise of appraisers (evaluators), in the field of teaching, seems to be especially important. Berliner (1987) reported on the results of a number of studies comparing the ratings of teacher effectiveness given by beginners (novices), advanced

EXHIBIT 1-5
OBSERVERS/EVALUATORS IN SREB STATES

STATE	NUMBER OBSERVERS	TYPE OBSERVERS	[illegible]	[illegible]	[illegible]	OBSERVERS AS EVALUATORS	[illegible]	EVALUATOR TITLE
Alabama	B 2 C 1	B) LEA option C) P,AP,S	B Yes C No	B LEA C LEA	B Supt./Prin C Prin	B Yes C Yes	B Videotape/Test C Videotape/Test	B No C No
Arkansas	B 1 C 1	B) P,AP C) P,AP,S	B No C No	B LEA C LEA	B Supt. C Supt.	B Yes C Yes	B Complete Training C No	B No C No
Florida	B 3 C 1	B) P,AP,S,IHE,T C) P	B Yes C No	B LEA/IHE C LEA	B Prin. C Prin.	B No C Yes	B Videotape/Test C No	B No C No
Georgia	B 3 C 1	B) P,T,Spec. C) P,AP,T,S	B Yes C No	B LEA/Regional Centers C LEA	B LEA Regional Centers C LEA	B No C Yes	B Videotape/Test/Field C Videotape/Test/Field	B Data Collector Assessment Specialist C No
Louisiana	B 3 C 3	B) P,T,IHE C) P,T,Ext	B Yes C Yes	B LEA C State/LEA	B No Policy C No Policy	B Yes C Yes	B Videotape/Test/Field C Videotape/Test/Field	B No C No
Mississippi	B 3 C 1	B) P,T,Ext. C) P,AP	B Yes C No	B State Pool C LEA	B Prin. C Prin.	B Yes C Yes	B Videotape/Test C Videotape/Test	B No C No
North Carolina	B 2 C 1 C 2 (Merit)	B) P, Mentor C) P,AP,T (merit)	B Yes C Yes (merit)	B LEA C LEA	B LEA C Prin.	B No C Yes	B No C Completing Training	B Mentor Observer/ Eval. C Observer/ Eval. (Teachers)
Oklahoma	B 3 C 1	B) P,T,IHE C) P	B Yes C No	B LEA/IHE C LEA	B Supt./Prin Teacher C Prin.	B Yes C Yes	B No C No	B Teacher Consul.)
Tennessee	B 1 C 1 C 3 (Career Ladder)	B) P C) T,P	B No C Yes (Career Ladder)	B LEA C LEA C State (Career Ladder)	B LEA C SDE/Prin.	B Yes C Yes	B LEA option/ complete Training C LEA option C Videotape (Career Ladder)	B No C Peer Eval. (Career Ladder)
Texas[1]	C 1	C) P	C No	C Other LEAs Contractors	C Co.	C Yes	C Videotape/Field	C No
Virginia[2]	B 3	B) Ext.	B No	B Former Educators	B LEA Recommends	B No	B Videotape/Test	B BTAP Recorder
W. Virginia	B 1 C 1	B) P,AP,S C) P,AP,S	B No C No	B LEA C LEA	B LEA C LEA	B LEA option C LEA option	B No C No	B No C No

Key: B = Beginning Teacher program; C = Continuing Teacher program
P = Principal; AP = Assistant Principal, S = Supervisor
T = Teacher, IHE = higher education person

[1] Texas has no state evaluation program for beginning teachers.

[2] Virginia has no state evaluation for continuing teachers.

FRENCH/SOUTHERN REGIONAL EDUCATION BOARD

BEST COPY AVAILABLE

beginners, and expert teachers. The studies indicate major differences in ratings of teacher effectiveness associated with the different levels of pedagogical expertise, including:

- Experts were able to monitor teachers and students more accurately than the advanced beginners and novices.
- Advanced beginners and novices reported considerable difficulty in comprehending teaching and learning events and behaviors.
- Different advanced beginners and novices reported contradictory interpretations of the same teaching and learning events and behaviors whereas different experts reported consistent events and behaviors.
- Advanced beginners and novices did not make evaluative comments on teaching and learning events and behaviors whereas experts combined interpretation with evaluation of the events and behaviors.
- Novices, advanced beginners, and experts differed in the assumptions, predictions, and hypothesis they made as they interpreted classroom events.
- Novices and advanced beginners demonstrated more variability in what they attended to and how they interpreted events than did the experts.

Berliner recommends that teachers be used as appraisers.

1.6.2 Appraiser Training

Training programs frequently are required in order to provide appraisers with the skills necessary for using an appraiser system, especially if the appraisal system is sophisticated or complex (Sims, Veres, and Heninger, 1987). Even with an effective appraisal system, dissatisfaction with the training aspect of the appraisal system can erase belief in the fairness of the entire appraisal process and cast doubts upon employee perceptions of organizational success (Daley, 1987).

Appraisers who receive the following types of training provide more accurate ratings than those who do not:

- guidelines on how to use the appraisal system, with opportunities to practice using the system, and with constructive feedback (Sims, Veres, and Heninger, 1987; Sims, 1988)
- job evaluation techniques (Hahn and Dipboye, 1988)
- the mechanics of appraisal, the specific requirements of the system, techniques for how to observe employee performance and to compare observations with established standards (Wehrenberg, 1988)
- rating standards and behavioral examples of rating dimensions (Daley, 1987)
- observation and decision-making skills (Hedge & Kavanaugh, 1988)

If appraisers do not possess substantial information about the jobs they are to appraise, then training also may become a critical mechanism for providing necessary job information. The importance of such job information was cited in the previous section (Hahn & Dipboye, 1988; Henenman, Wexely & Moore, 1987; Kamouri & Balzer, 1990; Smither & Reilly, 1989). Furthermore, appraisers who provide the most accurate and reliable ratings tend to be those who receive full information in the form of job descriptions and titles, as well as training (Hahn & Dipboye, 1988).

In teacher appraisal programs, initial and follow-up training (see exhibit 1-6) are key components. In the literature, training was found to range from a low of one to two days to a high of 20 days (French, Holdzkom & Kuligowksi, 1990). However, it was noted that "some programs may not be adequate to ensure accomplishment of the performance standards required of evaluators".

EXHIBIT 1-6

EVALUATOR TRAINING IN SREB STATES

[illegible]	TRAINING VARIABLES		
	[illegible]	[illegible]	FOLLOW-UP
Alabama	B 8 days C 8 days	B SDE/Turnkey C SDE/Turnkey	B Yes C Yes
Arkansas	B 1-2 days C 1-2 days	B SDE C SDE	B Yes C Yes
Florida	B 18 hours C LEA option	B SDE/Turnkey C LEA option	B Yes C LEA option
Georgia	B 50 hours C 50 hours	B SDE C SDE/Regional Assess. Ctr.	B Yes C No
Louisiana	B 60 hours C 60 hours	B SDE/Turnkey C SDE/Turnkey	B Yes C Yes
Mississippi	B 24-30 hours C 24-30 hours	B SDE C SDE	B No C No
North Carolina	B 30-80 hours C 84 hours	B SDE/Turnkey C SDE/Turnkey	B Yes C Optional
Oklahoma	B 2 days C 2 days	B SDE C SDE	B No C No
Tennessee	B LEA option C LEA option C 20 days (Career Ladder)	B SDE/Turnkey C SDE/Turnkey C SDE (Career Ladder)	B LEA option C LEA option C Yes (Career Ladder)
Texas	C 90 hours	C SDE/Turnkey	C Yes
Virginia	B 20 hours	B Developers	B No
W. Virginia	B LEA option C LEA option	B LEA C LEA	B LEA option C LEA option

FRENCH/SOUTHERN REGIONAL EDUCATION BOARD

BEST COPY AVAILAB

1.7 Summary of Literature and Research

1.7.1 Performance Appraisal in Business and Industry

The literature and research on performance appraisal in business and industry indicates increased interest and activity toward the development and validation of appraisal systems. The issues and problems addressed in the research are not dissimilar to those encountered in the development of performance appraisal systems for teachers. Overall, the causes for dissatisfaction with performance appraisals in business and industry are:

- lack of direct relationship of pay to performance
- unfairness or inequity in the process or results
- organizational problems
- bias
- politics

The most common measurement instruments in business and industry are scales based on trait-ratings, behavior, effectiveness, and hybrids of each of these. Some evidence was available, however, that the number of dimensions on rating scales was being reduced as it has been on the low inference instruments used for teacher appraisal.

The following recommendations are among the alternatives reported as ways to improve appraisal systems in business and industry:

- Document critical incidents and job behavior at the time of the behavior or shortly after the behavior occurs.
- Reduce the number of categories of performance to those that supervisors can reliability distinguish and clearly explain.
- Use group or team measures with white-collar workers.
- Use multiple sources and multiple raters.
- Give credit for good form, documented in specific critical incidents, rather than placing an undue emphasis on outcomes.

- Rate only extreme performance, excluding the middle.
- Define "good form" by citing specific positive and negative critical incidents.
- Use a coordination panel to process the evaluations.

1.7.2 Teacher Appraisals and Career Ladders

The research and literature on teacher appraisal systems and career ladders that seem to be enjoying some degree of success revealed the following common attributes:

- Performance observation instruments are designed to assess effective teaching as defined in the process-product research on teacher effectiveness.
- Observation records and other instruments are developed and validated through a systematic process involving practicing teachers and evaluation specialists.
- Reliability is established and reaffirmed through a regular process of extensive training, retraining and assessment of observers/evaluators.
- Multiple measures of performance are provided with more than one instrument and with two or more observations by two or more observers.
- All teachers are afforded access to all appraisal system guidelines, materials and observer training.
- Principals, assistant principals and peer teachers are the usual observers. Some systems use paid observers from outside the school or district.
- Evaluations are prepared immediately after the observations and other forms of data collection and feedback is provided to all teachers who are evaluated.
- Evaluation systems discriminate among levels of performance for various levels of the career ladder.
- In successful career ladder programs funding is available to reward all teachers who meet the required standards.
- Formative evaluation and professional development based on the knowledge and skills measured in the appraisal system is offered to all teachers.

The major unsolved issues' noted with teacher appraisal systems and career ladders in the literature and research were:

- Many current appraisal systems do not assess adequately the teacher's knowledge of content and ability to apply the knowledge to the range of learners and situations in the classroom.

- Most appraisal systems do not provide information on the relationship of teacher performance to student performance.

- There is little evidence that emerging research on higher order teaching skills has been included in teacher appraisal systems to differentiate between the good and best teachers.

- Incentive programs that focus on individual performance or evaluate the status of one teacher over another continue to meet with resistance.

- Programs that provide extra pay for extra work or focus on school-wide rather than individual incentives are increasingly popular.

BIBLIOGRAPHY

BIBLIOGRAPHY

Anonymous. (1988). "Assessing Pay-For-Performance Systems: Perception, Consistency Are the Keys to Effectiveness." *Employee Benefit Plan Review*, vol. 42, iss. 1, pp. 71-74.

Bannister, Brendan D.; Balkin, David B. (1990) "Performance Evaluation and Compensation Feedback Messages: An Integrated Model." *Journal of Occupational Psychology (UK)*, vol. 63, iss. 2, pp. 97-111.

Barrett, Gerald V.; Kernan, Mary C. (1987). "Performance Appraisal and Terminations: A Review of Court Decisions Since Brito V. Zia With Implications for Personnel Practices." *Personnel Psychology*, vol. 40, iss. 3, pp. 489-503.

Berliner, David C. (1988). "Implications of Studies of Expertise in Pedagogy for Teacher Education and Evaluation". *In New Direction for Teacher Assessment, Proceedings of 1988 ETS Invitational Conference*, Princeton.

Boyett, Joseph H.; Conn, Henry P. "Developing White-Collar Performance Measures". *National Productivity Review*, Summer, 1988, p. 209-218.

Brandt, Richard M. (1990). Incentive Pay and Career Ladders For Today's Teacher: A Study of Current Programs and Practices. State University of New York Press, New York.

Buford, James A. Jr.; Burkhalter, Betty B.; Jacobs, Grover T. (1988). "Link Job Descriptions to Performance Appraisals." *Personnel Journal*, vol. 67, iss. 6, pp. 132-140.

Buzzotta, V. R. (1988). "Improve Your Performance Appraisal." *Management Review*, vol. 77, iss. 8, pp. 40-43.

Buzzotta, V. R. (1989). "Improve Your Performance Appraisals." *Security Management*, vol. 33, iss. 6. pp. 94-100.

Campbell, Clifton P.; Armstrong, Richard B., Jr. (1988). "A Methodology for Testing for Job Task Performance: Part I." *Journal of European Industrial Training (UK)*, vol. 12, iss. 1, pp. 17-25.

Christensen, Margaret Howard. (1990). "Peer Auditing." *Nursing Management*, vol. 21, iss. 1, pp. 50-52.

Cook, Frederick W. (1989). "Making a Merit-Salary Program Work." *Journal of Compensation & Benefits*, vol. 5, iss. 1, pp. 45-50.

Cornell, Bonnie. (1986). "Appraising Grace: Evaluation." *Successful Executive (Canada)*, vol. 1, iss. 5, pp. 42-45.

Cornett, Lynn; Gaines, Gale (1991). "Linking Performance to Rewards for Teachers, Principals, and Schools". *The 1990 SREB Career Ladder Clearinghouse*. Southern Regional Education Board.

Denton, D. Keith. (1987). "Effective Appraisals: Key to Employee Motivation." *Industrial Engineering*, vol. 19, iss. 12, pp. 24-25, 46.

Edwards, Mark R. (1989). "Making Performance Appraisals Meaningful and Fair." *Business*, vol. 39, iss. 3, pp. 17-25.

Eyres, Patricia S. (1989). "Legally Defensible Performance Appraisal Systems." *Personnel Journal*, vol. 68, iss. 7, pp. 58-62.

Fox, William M. (1987/1988). "Improving Performance Appraisal Systems." *National Productivity Review*, vol. 7, iss. 1, pp. 20-27.

French, Russell L. , Holdzkom, David, Kuligowski, Barbara. "Teacher Evaluation In SREB States Stage I: Analysis and Comparison of Evaluation Systems.

Furtwengler, Carol B. (1987). " Lessons from Tennessee's Career Ladder Program". *Educational Leadership*, April 1987, pp. 66-69.

Gaugler, Barbara B.; Thornton, George C., III. (1989). "Number of Assessment Center Dimensions as a Determinant of Assessor Accuracy." *Journal of Applied Psychology*, vol. 74, iss. 4, pp. 611-618.

Geber, Beverely. (1988). "The Hidden Agenda of Performance Appraisals." *Training*, vol. 25, iss. 6, pp. 42-47.

Geis, A. Arthur. (1987). "Making Merit Pay Work." *Personnel*, vol. 64, iss. 1, pp. 52-60.

Geis, A. Arthur. (1989). "Making Merit Pay Work." *Canadian Manager (Canada)*, vol. 14, iss. 4, pp. 24-26.

Giffin, Margaret E. (1989). "Personnel Research on Testing, Selection, and Performance Appraisal." *Public Personnel Management,* vol. 18, iss. 2, pp. 127-137.

Goddard, Robert W. (1989). "Is Your Appraisal System Headed for Court?" *Personnel Journal* vol. 68, iss. 1, pp. 114-118.

Gomez-Mejia, Luis R. (1988). "Evaluating Employee Performance: Does the Appraisal Instrument Make a Difference?" *Journal of Organizational Behavior Management*, vol. 9, iss. 2, pp. 155-172.

Grant, Phillip C. (1988). "The Law of Employee Evaluations." *Management Solutions*, vol. 33, iss. 5, pp. 22-26.

Grosbeak, G. R. (1988) "Understanding People -- Physique & Personality." *Management Services (UK)*, vol. 32, iss. 5, pp. 18-23.

Hahn, David C.; Dipboye, Robert L. (1988). "Effects of Training and Information on the Accuracy of Reliability of Job Evaluation." *Journal of Applied Psychology*, vol. 73, iss. 2, pp. 146-153.

Hall, James L.; Posner, Barry Z.; Harder, Joseph W. (1989). "Performance Appraisal Systems: Matching Practice With Theory." *Group & Organization Studies*, vol. 14, iss. 1, pp. 51-69.

Heneman, Robert L.; Wexely, Kenneth N.; Moore, Michael L. (1987). "Performance-Rating Accuracy: A Critical Review." *Journal of Business Research*, vol. 15, iss. 5, pp. 431-448.

Hills, Frederick S.; Scott, K. Dow; Markham, Steven E.; Vest, Michael J. (1987). "Merit Pay: Just or Unjust Desserts." *Personnel Administrator*, vol. 32. iss. 9, pp. 52-59.

Hoevemeyer, Victoria A. (1989). "Performance-Based Compensation: Miracle or Waster?" *Personnel Journal*, vol. 68, iss. 7, pp. 64-68.

Jenkins, W. E.; McHarg. E. (1988). "Developing Professional Excellence." *Professional Safety*, vol. 33, iss. 5, pp. 17-20.

Kamouri, Anita L.; Balzer, William K. (1990). "The Effects of Performance Sampling Methods on Frequency Estimation, Probability Estimation, and Evaluation of Performance Information." *Organizational Behavior & Human Decision Process*, vol. 45, iss. 2, pp. 285-316.

Korinek, John T.; Thobe, Deborah J. (1989). "Salary Increases - How Much Is Enough." *Business Communications Review*, vol. 19, iss. 5, pp. 30-33.

Kuehn, Kathryn. (1988). "Appraisal Systems -- Designing Performance Appraisal Forms." *Credit Union Executive*, vol. 28, iss. 2, pp. 22-24.

Lawrie, John. (1989). "Steps Toward an Objective Appraisal." *Supervisory Management*, vol. 34, iss. 5, pp. 17-24.

Lee, Charles. (1989). "Poor Performance Appraisals Do More Harm Than Good." *Personnel Journal*, vol. 68, iss. 9, pp. 91-99.

Lee, Charles. (1990). "Smoothing Out Appraisal Systems." *HR Magazine*, vol. 35, iss. 3, pp. 72-76.

Liccione, William J. (1988). "Pay for Performance: What Are You Willing to Pay For?" *Compensation & Benefits Management*, vol. 3 iss. 3, pp. 251-253.

Locher, Alan H.; Teel, Kenneth S. (1988). "Appraisal Trends." *Personnel Journal*, vol. 67, iss. 9, pp. 139-145.

Lupton, Daniel E. (1989). "Two Steps to Make Performance Ratings More Accurate." *Journal of Compensation & Benefits*, vol. 5, iss. 3, pp. 191-192.

Marcoulides, George A.; Mills, R. Bryant. (1988). "Employee Performance Appraisals: A New Technique." *Review of Public Personnel Administration*, vol. 8, iss. 3, pp. 105-113.

Metz, Edmund J. (1988). "Designing Legally Defensible Performance Appraisal Systems." *Training & Development Journal*, vol. 42, iss. 7, pp. 47-51.

Moss, Stanley M.; Moen, Ronald D. (1989). "Appraise Your Performance Appraisal Process; The Performance Appraisal System; Deming's Deadly Disease." *Quality Progress*, vol. 22. iss. 11, pp. 58-66.

Murphy, Kevin R.; Constans, Joseph I. (1987). "Behavioral Anchors as a Source of Bias in Rating." *Journal of Applied Psychology*, vol. 72, iss. 4, pp. 573-577.

Nelson-Horchler, Joan. (1988). "Performance Appraisals." *Industry Week*, vol. 237, iss. 6, pp. 61-63.

Pelle, Denise; Greenhalgh, Leonard. (1987). "Developing the Performance Appraisal System." *Nursing Management*, vol. 18, iss. 12, pp. 37-44.

Pritchard, Robert D.; Roth, Phillip L.; Roth Patricia Galgay; Watson, Margaret D.; Jones, Steven D. (1989). "Incentive Systems: Success by Design." *Personnel*, vol. 66, iss. 5, pp. 63-68.

Pulhamus, Aaron R. (1989). "Performance Appraisal - Why and How You Should Implement a Due Process Approach to Discipline." *Supervision*, vol. 50, iss. 6.

Rollins, Thomas. (1987). "Pay for Performance: The Pros and Cons." *Personnel Journal*, vol. 66, iss. 6, pp. 104-111.

Schuster, Jay R.; Zinghelm, Patricia K. (1988). "Incentive Plans That Work." ABA Banking *Journal*, vol. 80, iss. 9, pp. 62-64.

Shelly, Paul N. (1987). "Translating A Performance Appraisal Into A Salary Adjustment." *Journal of Compensation & Benefits*, vol. 3, iss. 3, pp. 149-153.

Sims, Ronald R. (1988). "Training Supervisors in Employee Performance Appraisals." *Journal of European Industrial Training (UK)*, vol. 12, iss. 8, pp. 26-31.

Sims, Ronald R.; Veres, John G., III; Heninger, Susan M. (1987). "Training Appraisers: An Orientation Program for Improving Supervisory Performance Ratings." *Public Personnel Management*, vol. 16, iss. 1, pp. 37-46.

Smither, James W.; Barry, Stephen R.; Reilly, Richard R. (1989). "An Investigation of the Validity of Expert True Score Estimates in Appraisal Research." *Journal of Applied Psychology*, vol. 74, iss. 1, pp. 143-151.

Smither, James W.; Reilly, Richard R. (1989). "Relationship Between Job Knowledge and the Reliability or Conceptual Similarity Schemata." *Journal of Applied Psychology*, vol. 74, iss. 3, pp. 530-534.

Stufflebeam, David L. (1988). The Personnel Evaluation Standards, *Joint Committee on Standards for Educational Evaluation*, Sage Publications, Newberry Park.

Svatko, James E. (1989). "Simplifying the Performance Appraisal." *Small Business Reports*, vol. 14, iss. 3, pp. 30-33.

Thayer, Frederick C. (1987). "Performance Appraisal and Merit Pay Systems: The Disasters Multiply." *Review of Public Personnel Administration*, vol. 7, iss. 2, pp. 36-53.

Tziner, Aharon, Latham, Gary P. (1989). "The Effects of Appraisal Instrument, Feedback and Goal-Setting On Worker Satisfaction and Commitment." *Journal of Organizational Behavior (UK)*, vol. 10, iss. 2, pp. 145-153.

Vinton, Karen L. (1990). "Documentation That Gets Results." *Personnel*, vol. 67, iss. 2, pp. 46-46.

Webster, George D. (1988). "The Law of Employee Evaluations." *Association Management*, vol. 40, iss. 5, pp. 118-119.

Wehrenberg, Stephen B. (1988). "Train Supervisors to Measure and Evaluate Performance." *Personnel Journal*, vol. 57, iss. 2, pp. 77-81.

Weiss, Donald H. (1988). "The Legal Side of Performance Appraisals." *Management Solutions*, vol. 33, iss. 5, pp. 27-31.

701a/bib2

APPENDIX C
Opinion Survey and Instructions

EVALUATION OF THE TEXAS TEACHER APPRAISAL SYSTEM

INSTRUCTIONS FOR COMPLETION OF SURVEY

As required by the 71st Texas Legislature, the Texas Education Agency is conducting an evaluation of the Texas Teacher Appraisal System (TTAS). An important component of the evaluation is the analysis of data collected from a questionnaire distributed to a sample of teachers and administrators. The purpose of these efforts is to assess teachers' and administrators' perceptions of the impact of the TTAS upon the overall school operation and upon individuals.

Please complete the enclosed questionnaire (front and back) and return it in the envelope provided. No information is requested which will allow the identification of individual respondents. Completed questionnaires should be returned to the Texas Education Agency no later than:

November 16, 1990.

For your information, issues for this survey were generated from interviews with educators knowledgeable about the TTAS. In addition, representatives of the following professional organizations were interviewed: Association of Texas Professional Educators, Texas Association of Secondary School Principals, Texas Classroom Teachers Association, Texas Elementary Principals and Supervisors Association, Texas Federation of Teachers, and Texas State Teachers Association. Each of these organizations encourages you to participate in this important study. The results will be presented to the State Board of Education and the 72nd Texas Legislature. The final report will include a summary of the issues related to the TTAS which were found to be important to teachers and administrators, as well as any recommendations for modification of the TTAS which arise from the evaluation study.

If you have questions about this evaluation, please contact David Stamman or Cherry Kugle in the Division of Program Evaluation at 512/463-9524.

SURVEY OF SUPERINTENDENTS', PRINCIPALS', AND TEACHERS' IMPRESSIONS OF THE TEXAS TEACHER APPRAISAL SYSTEM (TTAS)

Demographic Information: Please check ( ) or provide information requested for each of the following questions.

1. Position: ___ Superintendent ___ Teacher ___ Principal ___ Other
2. If applicable, indicate grade level or subjects taught: ______________________
3. ___ Total Years of Teaching Experience
4. ___ Years of Teaching Experience in this District
5. ___ Years of Administrative Experience
6. Gender: ___ Male ___ Female
7. Ethnic Background:
   - ___ American Indian
   - ___ Asian/Pacific Islander
   - ___ Black (Non-Hispanic)
   - ___ Hispanic
   - ___ White (Non-Hispanic)

INSTRUCTIONS: For each of the following questions, please circle either "C" to indicate "completely", "A" to indicate "adequately", "I", to indicate "inadequately", or "N" to indicate "not at all". If you feel you do not have enough experience with the TTAS to give an opinion about some items, please circle "DK" to indicate "don't know".

8. In your district, to what extent does the TTAS actually accomplish each of the following objectives?	**8.**					
(a) Reflect actual teaching ability or competence	(a)	C	A	I	N	DK
(b) Improve classroom teaching methods	(b)	C	A	I	N	DK
(c) Improve unsatisfactory teaching skills	(c)	C	A	I	N	DK
(d) Increase student achievement	(d)	C	A	I	N	DK
(e) Identify ineffective teaching practices	(e)	C	A	I	N	DK
(f) Identify exceptional quality of teaching methods	(f)	C	A	I	N	DK
(g) Objectively and impartially appraise teacher performance	(g)	C	A	I	N	DK
(h) Result in more fairness of teacher appraisal than earlier evaluation practices	(h)	C	A	I	N	DK
(i) Improve morale of teachers	(i)	C	A	I	N	DK
(j) Standardize appraisal of teachers	(j)	C	A	I	N	DK
9. To what extent does the TTAS contain criteria and indicators which assess teaching behaviors for each of the following areas?	**9.**					
(a) Different ability levels of students	(a)	C	A	I	N	DK
(b) Basic teaching practices for most subjects	(b)	C	A	I	N	DK
(c) Teaching of critical thinking skills for students	(c)	C	A	I	N	DK
(d) Different grade levels	(d)	C	A	I	N	DK
SPECIALIZED SUBJECTS OR AREAS SUCH AS:						
(e) Art	(e)	C	A	I	N	DK
(f) Vocational education	(f)	C	A	I	N	DK
(g) Labs	(g)	C	A	I	N	DK
(h) Special education	(h)	C	A	I	N	DK
10. To what extent does the TTAS have the potential to identify differences between master and regular teachers?	**10.**	C	A	I	N	DK

BEST COPY AVAILABLE

11. To what extent are each of the following alternatives for appraisal of master teachers appropriate?	**11.**
(a) Use of present TTAS instrument	(a) C A I N DK
(b) Use of different scoring procedures with the same TTAS instrument	(b) C A I N DK
(c) Adding other domains to the present TTAS instrument	(c) C A I N DK
(d) Using a different instrument and procedure	(d) C A I N DK
12. To what extent does the TTAS clearly describe or define each of the following factors?	**12.**
(a) The criteria for appraising effective teaching practices	(a) C A I N DK
(b) The indicators for appraising effective teaching practices	(b) C A I N DK
(c) Observable teaching behaviors	(c) C A I N DK
(d) Quantifiable teaching behaviors	(d) C A I N DK
(e) Quality of teaching performance	(e) C A I N DK
13. To what extent are scores on TTAS influenced by factors other than teaching performance, such as the factors listed below?	**13.**
(a) Personal subjectivity or bias of the appraisers	(a) C A I N DK
(b) Time and effort required to provide evidence of "below expected standard" ratings	(b) C A I N DK
(c) Time and effort required to provide evidence of "exceptional quality" ratings	(c) C A I N DK
(d) Pressure from school or district officials to produce a certain number and type of ratings.	(d) C A I N DK
(e) Jargon and buzz words contained in statements or indicators	(e) C A I N DK
(f) Availability of funds	(f) C A I N DK
14. To what extent is it appropriate to award exceptional quality points in each of the following ways?	**14.**
(a) Award exceptional quality points for each indicator in a criterion	(a) C A I N DK
(b) Award exceptional quality points for the entire criterion in a holistic way	(b) C A I N DK
(c) Award exceptional quality points for each indicator and for the whole criterion	(c) C A I N DK
15. To what extent are the following features of the TTAS appropriate for the appraisal of effective teaching practices?	**15.**
(a) Use of pre-conferences	(a) C A I N DK
(b) Use of post-conferences	(b) C A I N DK
(c) Length of time periods of observation	(c) C A I N DK
(d) Frequency of observations	(d) C A I N DK
(e) Monitoring of appraiser procedures	(e) C A I N DK
(f) Quality of training and appraisers	(f) C A I N DK
(g) Frequency of training for appraisers	(g) C A I N DK

BEST COPY AVAILABLE

(CONTINUED) **15. To what extent are the following features of the TTAS appropriate for the appraisal of effective teaching practices?**	15. (Con't)					
(h) Consistency of interpretations of TTAS indicators by appraisers	(h)	C	A	I	N	DK
(i) Specific subject area knowledge of appraisers	(i)	C	A	I	N	DK
(j) Pedagogical knowledge and experience of appraisers	(j)	C	A	I	N	DK
(k) Capability of appraisers to use the TTAS accurately as it is presently designed	(k)	C	A	I	N	DK
(l) Ability of appraisers to discriminate polar extremes of effective and ineffective teaching practices	(l)	C	A	I	N	DK
(m) Ability of appraisers to discriminate the full range of teaching practices	(m)	C	A	I	N	DK
16. To what extent should the TTAS be linked to the following aspects of employment and rewards for teachers?	16.					
(a) Contract renewal or termination	(a)	C	A	I	N	DK
(b) Career Ladder decisions	(b)	C	A	I	N	DK
(c) Campus Assignment	(c)	C	A	I	N	DK
17. To what extent should student outcomes be included in the appraisal of teachers?	17.	C	A	I	N	DK

18. Which of the following statements most accurately reflects your opinion of the relationship between the TTAS and its implementation?

___ (a) The major difficulty is with the TTAS itself, not how it is implemented in my district.

___ (b) The major difficulty is with the way the TTAS is implemented, not with the TTAS itself.

___ (c) There are major difficulties with both the TTAS and the way it is implemented in my district.

___ (d) There are no major difficulties with either implementation or the TTAS itself.

19. Which of the following options regarding the career ladder would you prefer to see implemented? (Check One)

( ) Abolish
( ) Abolish and provide monetary incentives to campuses for high performance
( ) Retain, as is
( ) Retain and fully fund
( ) Increase minimum salaries in lieu of career ladder system
( ) None of the above

20. If you are a teacher, what level of the career level are you currently assigned? ____________

Recommendations for Improvement

If you believe that changes should be made to the current appraisal system, please list the changes below.

____
____
____
____
____

THANK YOU FOR YOUR TIME IN COMPLETING THIS QUESTIONNAIRE. PLEASE RETURN THIS QUESTIONNAIRE IN THE ENCLOSED POSTAGE PAID ENVELOPE TO THE TEXAS EDUCATION AGENCY.

701/TTASQUES.2

BEST COPY AVAILABLE

MGT of America, Inc.

P.O. Box 38430 • 2425 Torreya Dr. • Tallahassee, FL 32315 • (904) 386-3191 • FAX (904) 385-4501

October 2, 1990

Mr. John W. Sullivan
Superintendent
Red Oak Independent School District
P.O. Box 160
Red Oak, Texas 75154-0160

Dear Mr. Sullivan:

Thank you for agreeing to participate in the interviews MGT is conducting as part of an evaluation of the Texas Teacher Appraisal System (TTAS). The purpose of these interviews, as indicated in the September 24, 1990 letter that you received from Deputy Commissioner Lynn Moak of the Texas Education Agency, is to assess teachers' and administrators' perceptions of the impact of the TTAS upon the overall school operation and upon individuals.

Mr. Scott Gilber of MGT is scheduled to be in your office on October 9, 1990 at 8:15 a.m. to begin the interviews. He can begin with you or with the TTAS appraiser or the teachers, whichever is most convenient for you. Each interview will take approximately 45 minutes.

A copy of the interview questionnaire is enclosed. You may want to share a copy of it with the appraiser and teachers that you have selected prior to the interviews.

We look forward to our visit to your school district and the interviews with you and selected staff members.

Sincerely,

Garfield Wilson

Garfield W. Wilson
Project Director

Enclosure

701/ttaslt

APPENDIX D
Interview Protocol and List of Districts

INTERVIEW GUIDE
TEXAS TEACHER APPRAISAL SYSTEM (TTAS)

Introduction: We are conducting this interview as part of an evaluation of the entire Texas Teacher Appraisal system, including the TTAS instrument, the scoring procedures, and other parts of the system. Your answers to these interview questions are especially important since they will help us to clarify some of the results of the written survey conducted this month on the appraisal system. You also may have some issues that you would like to discuss which were not explored in the survey.

Your response to this interview is confidential. The information that you provide will only be reported as part of aggregated group responses. As part of our reporting on the interview group, we do need to record some basic demographic information.

Demographic Information

1. What is your position? __Supt __Prin __Tchr __Othr(__________)

2. What grade level and subjects do you teach, if any?__________________

3. How many years of total teaching experience do you have?________________

4. How many years of teaching experience do you have in this district?___________

5. How many years of administrative experience do you have?_____________

6. Gender: __Male __Female

7. Ethnic background: __Amer Indian __Asian/Pacific Islander

__Black (non-Hispanic) __Hispanic __White

Impact on Teaching Methods and Teachers

8. Does the TTAS actually accomplish each of the following objectives? If so, in what ways?

Y N a. reflect actual teaching ability or competence

Y N b. improve classroom teaching methods

Y N c. improve unsatisfactory teaching skills

Y N d. increase student achievement

Y N e. identify ineffective teaching practices

Y N f. identify exceptional quality of teaching methods

Y N g. objectively appraise teacher performance

Y N h. result in more fairness of teacher appraisal than earlier evaluation practices

Y N i. improve morale of teachers

Y N j. standardize appraisal of teachers

9. Does the TTAS contain criteria and indicators which assess teaching behaviors for each of the following areas? If so, in what ways?

Y N a. different ability levels of students

Y N b. basic teaching practices for most subjects

Y N c. teaching of critical thinking skills for students

Y N d. different grade levels

e. specialized subject areas such as:

Y N (1) art

Y N (2) vocational eduction

Y N (3) labs

Y N (4) special education

Y N (5) other(s)____

Y N 10. Does the TTAS have the potential to identify differences between master and regular teachers? If so, in what ways?

11. Are any of the following alternatives appropriate for appraisal of master teachers? If so, why? how?

Y N a. use of the present TTAS instrument

Y N b. use of different scoring procedures with the same TTAS instrument

Y N c. adding other domains to the present TTAS instrument

Y N d. using a different instrument and procedure

12. In your opinion, does the TTAS clearly describe or define each of the following factors? If not, what changes would you make?

Y N a. the criteria for appraising effective teaching practices

Y N b. the indicators for appraising effective teaching practices

Y N c. observable teaching behaviors

Y N d. quantifiable teaching behaviors

Y N e. quality of teaching performance

13. Do you believe that the TTAS scores are influenced by factors other than teaching
Y N performance? If so, what are the factors and how do they affect the scores?

(Possible prompts: subjectivity or bias of appraiser, time and effort in providing evidence for "below standard" and "exceptional quality" ratings, pressure from school or district officials, jargon or buzz words, availability of local funds)

14. How should exceptional quality points be awarded?

15. a. What parts of the TTAS do you believe are appropriate for a system to appraise effective teaching practices? Why?

(Possible prompts: pre or post-conferences, observation time periods and frequency, monitoring or training of appraisers, consistency of interpretation of TTAS indicators by appraisers, background of appraisers, capability of appraisers to use TTAS accurately as it is presently designed, ability to discriminate polar extremes and the full range of teaching practices.)

b. What parts of the TTAS do you believe are NOT appropriate for a system to appraise effective teaching practices? Why not?

(Possible prompts: same as above.)

16. Do you believe that the TTAS should be linked to the following aspects of employment and rewards for teachers? If so, why? If not, why not?

Y N a. contract renewal or termination

Y N b. career ladder decisions

Y N c. campus assignment

17. To what extent should student outcomes be included in the appraisal of teachers?

18. What are the major difficulties with the TTAS and with how it is implemented in your district?

(Possible prompts: are the major difficulties with the TTAS or with how it is implemented or with both or with neither?)

19. What changes do you believe should be made to improve the TTAS?

(Possible prompts: abolish, retain as is, abolish and fund campus incentive programs, retain and fully fund, increase minimum salaries in lieu of career ladder, other options.)

Thank you for your time in answering these questions. Your answers will help the TEA to make improvements in the appraisal system. If you think of anything else you would like to add, you can contact me at this number and address. (Provide business card.)

Districts Selected for Interviews

031901	Brownsville	1
125901	Alice	2
107901	Athens	7
070911	Red Oak	10
220905	Fort Worth	11
028902	Lockhart	13
025902	Brownwood	15
188903	Highland Park	16
071905	Ysleta	19
163904	Hondo	20

APPENDIX E
Analyze Categories for TTAS and Career Ladder

TEXAS EDUCATION AGENCY
MEAN TOTAL APPRAISAL SCORES FOR SPRING, 1989 AND SPRING, 1990

NBR DIST	CATEGORY	TOTAL APPRAISAL SCORE SPRING 89	TOTAL APPRAISAL SCORE SPRING 90
ADA GROUPINGS			
6	OVER 50,000	153.5	158.4
15	25,000 TO 49,999	157.7	160.4
44	10,000 TO 24,999	151.0	158.7
41	5,000 TO 9,999	152.9	156.0
92	3,000 TO 4,999	152.0	157.0
115	1,600 TO 2,999	152.3	156.1
125	1,000 TO 1,599	150.8	155.0
207	500 TO 999	150.3	153.8
409	UNDER 500	147.8	152.6
DISTRICT TYPE			
8	MAJOR URBAN	154.8	158.3
61	MAJOR SUBURBAN	154.9	160.1
23	OTHER CENTRAL CITY	150.0	157.8
74	OTHER CC SUBURBAN	152.6	156.0
65	INDEPENDENT TOWN	152.6	156.1
82	NON-METRO FAST GROWING	151.1	155.5
271	NON-METRO STABLE	150.7	155.6
470	RURAL	148.2	152.9
WEALTH (MEDIAN=$157,104)			
105	UNDER $82,513	148.9	154.4
105	$82,513 TO $99,523	155.5	156.2
105	$99,524 TO $115,679	151.2	155.2
106	$115,680 TO $133,839	151.5	155.8
105	$133,840 TO $157,103	154.4	157.8
105	$157,104 TO $181,840	152.4	157.1
106	$181,841 TO $220,234	152.9	159.3
105	$220,235 TO $287,975	155.7	159.9
105	$287,976 TO $443,845	151.6	157.8
101	OVER $443,845	149.8	158.6
6	SPECIAL DISTRICTS	154.2	158.6
WEALTH (ST AVG=$205,696)			
698	UNDER $205,696	152.5	156.5
350	OVER $205,696	153.0	159.0
6	SPECIAL DISTRICTS	154.2	158.6
WEALTH BY EQUAL RADA PER GROUP			
25	UNDER $51,611	150.1	153.9
70	$51,611 TO < $79,599	147.6	152.8
75	$79,599 TO < $93,431	156.6	159.0
113	$93,431 TO < $111,174	150.9	154.6
82	$111,174 TO < $124,674	152.8	156.5
56	$124,674 TO < $133,840	150.2	155.1
63	$133,840 TO < $147,075	151.3	154.7
40	$147,075 TO < $156,339	157.4	160.7
32	$156,339 TO < $163,010	152.4	157.1
39	$163,010 TO < $174,981	152.3	157.1
58	$174,981 TO < $190,253	155.6	159.7
49	$190,253 TO < $206,247	153.2	157.7
29	$206,247 TO < $217,719	150.1	160.5
48	$217,719 TO < $244,860	152.2	157.2
41	$244,860 TO < $268,755	151.5	158.6
1	$268,755 TO < $269,138	158.9	160.0
33	$269,138 TO < $298,578	157.8	163.1
53	$298,578 TO < $367,048	152.8	156.8
13	$367,048 TO < $388,497	146.3	156.4
128	$388,497 AND OVER	153.5	159.6
6	SPECIAL DISTRICTS	154.2	158.6
TOTAL TAX EFFORT (ST AVG=$0.9606)			
260	UNDER 0.8092	151.0	154.6
262	0.8092 TO UNDER 0.9716	152.4	158.3
263	0.9716 TO UNDER 1.1289	153.4	157.7
263	1.1289 AND OVER	153.5	157.6
6	SPECIAL DISTRICTS	154.2	158.6
M&O EFF. TAX EFFORT (MEDIAN=$0.8130)			
260	UNDER 0.6780	152.1	155.9
263	0.6780 TO 0.8129	151.8	157.7
262	0.8130 TO 0.9647	153.6	158.4
263	OVER 0.9647	153.8	157.5
6	SPECIAL DISTRICTS	154.2	158.6
M&O EFF. TAX EFFORT (ST AVG=$0.8150)			
528	UNDER 0.8150	151.8	157.1
520	OVER 0.8150	153.8	158.0
6	SPECIAL DISTRICTS	154.2	158.6
1,054	STATE TOTAL	152.7	157.5

TEXAS EDUCATION AGENCY
MEAN TOTAL APPRAISAL SCORES FOR SPRING, 1989 AND SPRING, 1990

NBR DIST	CATEGORY	TOTAL APPRAISAL SCORE SPRING 89	TOTAL APPRAISAL SCORE SPRING 90
APR GROUPS: ENROLL ¦WEALTH¦% LOW INC			
190	<1K ¦ < AVG. ¦ <40%	148.5	152.7
190	<1K ¦ < AVG. ¦ >=40	149.6	153.5
123	<1K ¦ > AVG. ¦ <40%	148.8	153.4
96	<1K ¦ > AVG. ¦ >=40	149.1	153.2
90	1K TO < 3K ¦ < AVG. ¦ <40%	152.8	156.4
85	1K TO < 3K ¦ < AVG. ¦ >=40	149.0	153.5
41	1K TO < 3K ¦ > AVG. ¦ <40%	152.4	156.1
25	1K TO < 3K ¦ > AVG. ¦ >=40	153.1	156.2
64	3K TO < 10K ¦ < AVG. ¦ <40%	154.9	158.2
37	3K TO < 10K ¦ < AVG. ¦ >=40	149.3	153.3
35	3K TO < 10K ¦ > AVG. ¦ <40%	151.4	158.0
4	3K TO < 10K ¦ > AVG. ¦ >=40	153.9	155.6
22	>10K ¦ < AVG. ¦ <40%	153.3	158.0
20	>10K ¦ < AVG. ¦ >=40	154.1	157.9
19	>10K ¦ > AVG. ¦ <40%	154.0	161.7
7	>10K ¦ > AVG. ¦ >=40	153.5	159.0
6	SPECIAL DISTRICTS	154.2	158.6
SPTB HIGHEST CATEGORY			
325	RESIDENTIAL	153.6	158.6
332	LAND	150.0	153.4
186	OIL AND GAS	149.2	155.1
202	BUSINESS	152.2	156.8
6	SPECIAL DISTRICTS	154.2	158.6
SMALL/SPARSE ADJSTMNT (ST AVG=29.7%)			
290	NO SMALL/SPARSE ADJUSTMENT	153.3	158.2
192	UNDER 22.0%	150.7	154.9
191	22.0% TO UNDER 31.4%	150.2	153.3
192	31.4% TO UNDER 36.8%	147.0	152.7
189	36.8% AND OVER	149.1	153.4
PDI LEVEL (MEDIAN=1.07)			
265	UNDER 1.05	149.8	154.0
122	1.05 TO UNDER 1.06	149.8	154.4
245	1.06 TO UNDER 1.08	150.2	156.2
243	1.08 TO 1.11	152.3	156.4
179	1.11 AND OVER	153.5	158.4
OPERATING COST/RADA (ST AVG= $3,837)			
211	UNDER $3,476	154.9	157.9
212	$3.476 TO $3,806	151.0	156.8
212	$3,807 TO $4,211	154.4	158.4
212	$4,212 TO $4,940	150.6	157.4
207	OVER $4,940	152.5	156.6
ESC REGION			
38	I EDINBURG	150.8	154.6
44	II CORPUS CHRISTI	149.1	154.2
41	III VICTORIA	155.1	158.1
55	IV HOUSTON	155.4	158.3
29	V BEAUMONT	155.2	159.0
57	VI HUNTSVILLE	151.6	155.8
98	VII KILGORE	143.5	154.7
48	VIII MT PLEASANT	152.3	158.0
40	IX WICHITA FALLS	141.8	157.1
79	X RICHARDSON	150.7	158.0
77	XI FORT WORTH	158.5	162.1
78	XII WACO	152.2	156.3
56	XIII AUSTIN	154.1	158.1
47	XIV ABILENE	150.1	155.1
45	XV SAN ANGELO	146.4	150.8
68	XVI AMARILLO	149.9	155.4
59	XVII LUBBOCK	156.8	161.4
33	XVIII MIDLAND	152.3	155.1
12	XIX EL PASO	161.5	161.5
50	XX SAN ANTONIO	150.4	155.8
TEAMS: PCT PASSING ALL TESTS TAKEN			
78	UNDER 60%	145.6	150.8
367	60% TO UNDER 75%	153.1	157.1
546	75% TO UNDER 90%	152.8	158.5
63	90% AND OVER	155.9	159.1
DENSITY (ST AVG=11.34 RADA/SQ MI)			
564	LESS THAN 5	150.6	154.5
277	5 TO UNDER 20	150.1	154.9
116	20 TO UNDER 100	151.5	156.9
91	100 AND OVER	154.3	159.1
6	SPECIAL DISTRICTS	154.2	158.6
1,054	STATE TOTAL	152.7	157.5

TEXAS EDUCATION AGENCY

MEAN TOTAL APPRAISAL SCORES FOR SPRING, 1989 AND SPRING, 1990

NBR DIST	CATEGORY	TOTAL APPRAISAL SCORE SPRING 89	TOTAL APPRAISAL SCORE SPRING 90
RADA CHG:88/89-89/90 (ST AVG=1.57%)			
461	DECLINING RADA	148.7	155.3
311	0% TO UNDER 3%	155.1	158.6
166	3% TO UNDER 6%	154.1	158.6
68	6% TO UNDER 10%	150.4	155.6
48	10% AND OVER	152.5	163.2
PCT BLACK PUPILS (ST AVG=14.50%)			
623	UNDER 5%	152.2	156.1
136	5% TO UNDER 10%	153.0	158.2
149	10% TO UNDER 20%	155.6	159.9
73	20% TO UNDER 30%	151.1	156.6
61	30% TO UNDER 50%	150.8	157.4
12	50% AND OVER	151.6	157.5
PCT HISPANIC PUPILS (ST AVG=33.10%)			
310	UNDER 5%	151.5	156.4
170	5% TO UNDER 10%	157.0	161.1
170	10% TO UNDER 20%	149.8	157.5
101	20% TO UNDER 30%	153.8	158.0
125	30% TO UNDER 50%	152.6	157.3
178	50% AND OVER	151.7	155.4
PCT MINORITY PUPILS (ST AVG=49.70%)			
112	UNDER 5%	152.6	156.7
123	5% TO UNDER 10%	152.7	156.0
202	10% TO UNDER 20%	154.4	158.3
143	20% TO UNDER 30%	155.7	160.9
228	30% TO UNDER 50%	150.4	157.0
246	50% AND OVER	152.6	156.8
PERCENT LOW INCOME (ST AVG=39.75%)			
151	UNDER 20%	156.7	161.2
212	20% TO UNDER 30%	152.8	156.9
225	30% TO UNDER 40%	149.3	155.9
293	40% TO UNDER 60%	153.9	157.7
118	60% TO UNDER 80%	149.3	155.7
55	80% AND OVER	150.1	153.8
AVG. TEACHER EXPER (ST AVG=11.3 YRS)			
266	UNDER 9.7 YEARS	150.6	155.4
260	9.7 TO UNDER 11.1 YEARS	154.5	158.2
260	11.1 TO UNDER 12.3 YEARS	154.0	158.3
268	12.3 YEARS AND OVER	150.2	156.9
AVG. TEACHER SALARY (ST AVG=$25,897)			
263	UNDER $23,176	149.7	152.6
265	$23,176 TO UNDER $24,237	151.0	155.6
265	$24,237 TO UNDER $25,425	151.1	156.7
261	$25,425 AND OVER	154.0	158.7
PCT MINORITY TCHRS (ST AVG=22.40%)			
591	UNDER 5%	155.1	159.1
186	5% TO UNDER 10%	151.2	157.2
127	10% TO UNDER 20%	152.9	157.1
44	20% TO UNDER 30%	152.6	156.9
48	30% TO UNDER 50%	149.3	157.2
58	50% AND OVER	153.2	155.9
% TCHRS W ADV DEGREE (ST AVG=32.0%)			
263	UNDER 19.0%	149.7	154.3
264	19.0% TO UNDER 26.9%	153.5	156.4
265	26.9% TO UNDER 34.9%	153.5	158.4
262	34.9% AND OVER	152.4	158.1
1,054	STATE TOTAL	152.7	157.5

T.E.A.

CAREER LADDER LEVELS FROM SPRING 89/90 13:34 MONDAY, FEBRUARY 11, 1991 1

NBR DIST	CATEGORY	PERCENT OF TEACHERS ON LEVEL ONE	PERCENT OF TEACHERS ON LEVEL TWO	PERCENT OF TEACHERS ON LEVEL THREE
ADA GROUPINGS				
6	OVER 50,000	30.9	41.9	27.2
15	25,000 TO 49,999	40.9	32.3	26.8
44	10,000 TO 24,999	43.9	37.6	18.5
41	5,000 TO 9,999	41.1	42.4	16.5
92	3,000 TO 4,999	40.9	45.8	13.2
115	1,600 TO 2,999	45.3	43.8	10.8
125	1,000 TO 1,599	45.9	45.4	8.7
208	500 TO 999	48.0	45.8	6.2
411	UNDER 500	53.1	40.6	6.2
DISTRICT TYPE				
8	MAJOR URBAN	33.4	39.2	27.4
61	MAJOR SUBURBAN	42.0	34.7	23.4
23	OTHER CENTRAL CITY	42.7	40.3	17.0
74	OTHER CC SUBURBAN	42.5	44.9	12.6
65	INDEPENDENT TOWN	40.3	45.1	14.6
82	NON-METRO FAST GROWING	50.2	40.2	9.6
271	NON-METRO STABLE	45.2	45.8	9.1
473	RURAL	51.1	42.6	6.2
WEALTH (MEDIAN=$157,104)				
105	UNDER $82,513	47.8	39.0	13.1
105	$82,513 TO $99,523	43.3	40.2	16.5
105	$99,524 TO $115,679	47.6	43.4	9.1
106	$115,680 TO $133,839	42.0	42.8	15.2
105	$133,840 TO $157,103	41.9	39.9	18.2
105	$157,104 TO $181,840	45.6	42.1	12.3
106	$181,841 TO $220,234	41.8	37.3	20.8
105	$220,235 TO $287,975	34.9	43.4	21.8
105	$287,976 TO $443,845	38.1	37.0	24.9
104	OVER $443,845	42.5	42.3	15.2
6	SPECIAL DISTRICTS	46.6	32.9	20.5
WEALTH (ST AVG=$205,696)				
698	UNDER $205,696	44.4	40.5	15.1
353	OVER $205,696	37.6	40.1	22.3
6	SPECIAL DISTRICTS	46.6	32.9	20.5
WEALTH BY EQUAL RADA PER GROUP				
25	UNDER $51,611	45.4	39.7	14.9
70	$51,611 TO < $79,599	49.3	38.4	12.3
75	$79,599 TO < $93,431	45.3	37.8	16.9
113	$93,431 TO < $111,174	44.9	45.0	10.1
82	$111,174 TO < $124,674	46.4	37.5	16.1
56	$124,674 TO < $133,840	39.3	47.6	13.1
63	$133,840 TO < $147,075	45.5	40.7	13.8
40	$147,075 TO < $156,339	39.0	37.6	23.4
32	$156,339 TO < $163,010	44.8	44.2	11.0
39	$163,010 TO < $174,981	44.8	40.4	14.8
58	$174,981 TO < $190,253	44.1	39.4	16.5
49	$190,253 TO < $206,247	43.1	37.4	19.5
29	$206,247 TO < $217,719	40.6	37.0	22.5
48	$217,719 TO < $244,860	43.9	42.4	13.7
41	$244,860 TO < $268,755	44.6	38.3	17.1
1	$268,755 TO < $269,138	0.0	69.0	31.0
33	$269,138 TO < $298,578	39.3	29.6	31.1
53	$298,578 TO < $367,048	39.3	42.3	18.4
13	$367,048 TO < $388,497	37.7	31.5	30.8
131	$388,497 AND OVER	40.4	43.2	16.5
6	SPECIAL DISTRICTS	46.6	32.9	20.5
TOTAL TAX EFFORT (ST AVG=$0.9606)				
262	UNDER 0.8092	44.6	42.8	12.6
263	0.8092 TO UNDER 0.9716	39.3	38.7	22.0
263	0.9716 TO UNDER 1.1289	42.4	42.0	15.7
263	1.1289 AND OVER	43.3	39.5	17.2
6	SPECIAL DISTRICTS	46.6	32.9	20.5
M&O EFF. TAX EFFORT (MEDIAN=$0.8130)				
262	UNDER 0.6780	45.4	40.5	14.2
263	0.6780 TO 0.8129	39.0	39.5	21.5
263	0.8130 TO 0.9647	42.8	40.9	16.2
263	OVER 0.9647	41.5	40.9	17.6
6	SPECIAL DISTRICTS	46.6	32.9	20.5
M&O EFF. TAX EFFORT (ST AVG=$0.8150)				
530	UNDER 0.8150	41.2	39.8	19.0
521	OVER 0.8150	42.3	40.9	16.8
6	SPECIAL DISTRICTS	46.6	32.9	20.5
1,057	STATE TOTAL	41.7	40.3	18.0

NBR DIST	APR GROUPS: ENROLL¦WEALTH¦% LOW INC	PERCENT OF TEACHERS ON LEVEL ONE	PERCENT OF TEACHERS ON LEVEL TWO	PERCENT OF TEACHERS ON LEVEL THREE
190	<1K ¦ < AVG. ¦ <40%	50.5	44.0	5.5

CAREER LADDER LEVELS FROM SPRING 89/90 13:34 MONDAY, FEBRUARY 11, 1991 2

NBR DIST	CATEGORY			PERCENT OF TEACHERS ON LEVEL ONE	PERCENT OF TEACHERS ON LEVEL TWO	PERCENT OF TEACHERS ON LEVEL THREE
190	<1K	< AVG.	>=40	51.8	42.0	6.2
124	<1K	> AVG.	<40%	47.0	45.2	7.8
98	<1K	> AVG.	>=40	50.6	43.2	6.2
90	1K TO < 3K	< AVG.	<40%	44.8	44.7	10.5
85	1K TO < 3K	< AVG.	>=40	47.1	44.3	8.6
41	1K TO < 3K	> AVG.	<40%	44.8	46.0	9.1
25	1K TO < 3K	> AVG.	>=40	46.2	46.7	7.1
64	3K TO < 10K	< AVG.	<40%	41.7	44.3	13.9
37	3K TO < 10K	< AVG.	>=40	44.1	43.0	12.9
35	3K TO < 10K	> AVG.	<40%	37.7	45.5	16.8
4	3K TO < 10K	> AVG.	>=40	37.9	44.2	18.0
22	>10K	< AVG.	<40%	43.0	37.8	19.2
20	>10K	< AVG.	>=40	43.3	36.3	20.4
19	>10K	> AVG.	<40%	39.6	32.9	27.5
7	>10K	> AVG.	>=40	30.6	42.3	27.1
6	SPECIAL DISTRICTS			46.6	32.9	20.5
	SPTB HIGHEST CATEGORY					
328	RESIDENTIAL			42.6	38.0	19.4
332	LAND			50.8	42.0	7.2
189	OIL AND GAS			43.4	47.7	9.0
202	BUSINESS			37.7	43.2	19.1
6	SPECIAL DISTRICTS			46.6	32.9	20.5
	SMALL/SPARSE ADJSTMNT (ST AVG=29.7%)					
290	NO SMALL/SPARSE ADJUSTMENT			40.3	39.5	20.1
192	UNDER 22.0%			45.8	45.4	8.9
192	22.0% TO UNDER 31.4%			50.0	44.3	5.7
192	31.4% TO UNDER 36.8%			53.2	39.9	6.8
191	36.8% AND OVER			50.3	43.8	5.8
	PDI LEVEL (MEDIAN=1.07)					
265	UNDER 1.05			49.7	43.8	6.5
124	1.05 TO UNDER 1.06			45.4	46.9	7.7
246	1.06 TO UNDER 1.08			45.0	44.4	10.6
243	1.08 TO 1.11			43.8	44.1	12.1
179	1.11 AND OVER			40.1	38.1	21.8
	OPERATING COST/RADA (ST AVG= $3,837)					
211	UNDER $3,476			43.9	38.9	17.2
212	$3,476 TO $3,806			43.8	39.7	16.5
212	$3,807 TO $4,211			37.4	43.2	19.4
212	$4,212 TO $4,940			41.1	38.0	20.9
210	OVER $4,940			42.5	45.0	12.5
	ESC REGION					
38	I EDINBURG			43.0	43.3	13.6
44	II CORPUS CHRISTI			48.2	40.8	11.0
41	III VICTORIA			38.4	46.2	15.5
55	IV HOUSTON			36.2	40.9	22.9
29	V BEAUMONT			44.0	42.7	13.3
57	VI HUNTSVILLE			47.3	43.1	9.6
98	VII KILGORE			39.0	48.5	12.5
48	VIII MT PLEASANT			44.3	47.6	8.1
40	IX WICHITA FALLS			45.0	51.8	3.3
79	X RICHARDSON			40.3	36.8	22.9
77	XI FORT WORTH			45.4	34.5	20.1
78	XII WACO			46.8	38.2	15.1
56	XIII AUSTIN			38.2	45.4	16.4
47	XIV ABILENE			42.5	42.7	14.8
45	XV SAN ANGELO			45.6	46.0	8.5
69	XVI AMARILLO			49.2	43.8	7.0
61	XVII LUBBOCK			41.4	38.9	19.8
33	XVIII MIDLAND			48.9	43.8	7.3
12	XIX EL PASO			45.5	26.3	28.3
50	XX SAN ANTONIO			40.0	37.7	22.3
	TEAMS: PCT PASSING ALL TESTS TAKEN					
79	UNDER 60%			49.9	38.6	11.4
368	60% TO UNDER 75%			40.7	40.4	18.9
547	75% TO UNDER 90%			42.3	40.3	17.4
63	90% AND OVER			37.0	42.3	20.7
	DENSITY (ST AVG=11.34 RADA/SQ MI)					
566	LESS THAN 5			47.5	45.5	7.0
278	5 TO UNDER 20			44.7	43.2	12.0
116	20 TO UNDER 100			44.3	42.7	13.0
91	100 AND OVER			38.7	37.6	23.7
6	SPECIAL DISTRICTS			46.6	32.9	20.5
1,057	STATE TOTAL			41.7	40.3	18.0
	RADA CHG:88/89-89/90 (ST AVG=1.57%)					
462	DECLINING RADA			43.8	41.5	14.7
312	0% TO UNDER 3%			39.0	40.2	20.8
166	3% TO UNDER 6%			42.4	38.9	18.7
68	6% TO UNDER 10%			48.2	39.0	12.8

NBR DIST	CATEGORY	PERCENT OF TEACHERS ON LEVEL ONE	PERCENT OF TEACHERS ON LEVEL TWO	PERCENT OF TEACHERS ON LEVEL THREE
49	10% AND OVER	49.2	41.9	8.8
PCT BLACK PUPILS (ST AVG=14.50%)				
626	UNDER 5%	44.6	40.6	14.7
136	5% TO UNDER 10%	42.7	37.7	19.6
149	10% TO UNDER 20%	39.2	40.4	20.4
73	20% TO UNDER 30%	45.6	43.4	11.0
61	30% TO UNDER 50%	34.7	41.6	23.7
12	50% AND OVER	47.1	39.4	13.5
PCT HISPANIC PUPILS (ST AVG=33.10%)				
311	UNDER 5%	42.5	45.7	11.8
170	5% TO UNDER 10%	43.1	37.6	19.4
170	10% TO UNDER 20%	43.7	40.6	15.6
101	20% TO UNDER 30%	43.7	36.6	19.7
127	30% TO UNDER 50%	34.2	42.6	23.2
178	50% AND OVER	45.1	39.0	15.9
PCT MINORITY PUPILS (ST AVG=49.70%)				
113	UNDER 5%	46.8	43.4	9.8
123	5% TO UNDER 10%	43.7	45.3	11.0
202	10% TO UNDER 20%	43.3	40.2	16.4
143	20% TO UNDER 30%	40.5	39.3	20.1
230	30% TO UNDER 50%	42.6	40.9	16.5
246	50% AND OVER	40.7	39.6	19.7
PERCENT LOW INCOME (ST AVG=39.75%)				
152	UNDER 20%	41.5	37.4	21.1
212	20% TO UNDER 30%	42.9	41.5	15.6
225	30% TO UNDER 40%	42.8	41.8	15.4
295	40% TO UNDER 60%	38.7	40.7	20.6
118	60% TO UNDER 80%	46.5	39.7	13.8
55	80% AND OVER	42.7	42.2	15.1
AVG. TEACHER EXPER (ST AVG=11.3 YRS)				
267	UNDER 9.7 YEARS	51.3	36.4	12.3
260	9.7 TO UNDER 11.1 YEARS	43.9	39.0	17.1
260	11.1 TO UNDER 12.3 YEARS	37.8	41.5	20.7
270	12.3 YEARS AND OVER	39.4	42.3	18.2
AVG. TEACHER SALARY (ST AVG=$25,897)				
264	UNDER $23,176	53.7	38.9	7.4
265	$23,176 TO UNDER $24,237	45.8	43.0	11.2
265	$24,237 TO UNDER $25,425	44.7	41.4	13.9
263	$25,425 AND OVER	38.6	39.4	22.1
PCT MINORITY TCHRS (ST AVG=22.40%)				
594	UNDER 5%	43.1	40.4	16.5
186	5% TO UNDER 10%	44.6	41.4	13.9
127	10% TO UNDER 20%	41.8	39.2	19.0
44	20% TO UNDER 30%	41.8	40.4	17.8
48	30% TO UNDER 50%	44.1	32.4	23.4
58	50% AND OVER	31.3	50.4	18.3
% TCHRS W ADV DEGREE (ST AVG=32.0%)				
264	UNDER 19.0%	49.5	40.4	10.1
265	19.0% TO UNDER 26.9%	46.1	39.6	14.2
265	26.9% TO UNDER 34.9%	44.1	38.1	17.9
263	34.9% AND OVER	35.3	42.7	22.0
1,057	STATE TOTAL	41.7	40.3	18.0

BEST COPY AVAILABLE

APPENDIX F
Recent Changes in TTAS and Comment from Assistant Commissioner for Professional Development

M E M O R A N D U M

February 25, 1991

To: Division of Program Evaluation

From: Richard Swain
Assistant Commissioner for Professional Development

Re: Evaluation Study of the Texas Teacher Appraisal System

After reviewing the draft of the evaluation study of the TTAS and discussing with the Professional Development staff the formulation of a collective response, the following observations are offered:

The Executive Summary of the report merits special attention. Frequently such summaries are too hastily written and do not reflect complexities that are often part and parcel of evaluation studies. By contrast, this executive summary is a particularly well written, comprehensive synopsis of the whole report. The summary appropriately underscores the primary focus of the report and fairly represents the complexities of the TTAS study. However, some observations about some of the study's particular conclusions and format are in order.

At best, this study could only examine "perceptions" of local control versus state control. Because of the nature of the SBOE Rule, certain issues of local option and flexibility are stipulated in the rule. While it may be true that teachers and administrators have different perceptions of the degree to which possibilities for local control are in reality implemented, the SBOE clearly defines a number of issues -- such as stricter performance criteria, the possibility of locally generated funds supplementing state career ladder allotments, etc. -- that presuppose local control options.

An objectivity/subjectivity tension may be a significant factor which offers some explanation for administrators' opposition to applying EQs at the indicator level. Such a change could require a higher degree of instructional expertise to articulate a realistically operational strategy for assigning EQ points -- an expressly subjective determination in what is cast as an essentially objective assessment exercise.

Although it is possible for a case to be made for linking rewards for high performing campuses on the Academic Excellence Indicator System instead of retaining a form of the career ladder, such a case is not expressly supported by data gathered in this study. While it is possible to argue that extrapolating such an argument may be done, the data gathered and presented in this study do not explicitly support the case.

Data gathered from a major TTAS survey conducted by the Division of Management Assistance and Personnel Development (MAPD) in 1989 identified seven areas of improvement in the Texas Teacher Appraisal

BEST COPY AVAILABLE

System. Those improvement areas were further isolated through extensive discussions with teachers, practicing appraisers, ESC and college/university TTAS lead trainers, and leaders of teacher and administrator organizations. The areas of improvement are as follows:

Assessment of Exceptional Quality. Advisory group members and respondents indicated that teachers and appraisers need additional clarification about what exceptional quality in teaching means and entails;

Inter-Rater Reliability. Eighteen (18) performance indicators in the TTAS instrument were identified by respondents as problematic and difficult to evaluate because of their high subjectivity. For example, lf- instruction at appropriate level, 4c-appropriate pace, and 6a-appropriate introduction;

Quality of TTAS Certification and Update Training. Advisory group members and respondents indicated that the quality of the videotape proficiencies need to be improved and that the written examination needed to be changed from a recall-type test to an application-type test;

Delivery of Information and Training to Teachers. Advisory group members indicated that teachers need to be better informed about TTAS practices and procedures and where appropriate, they need to be involved in developing them;

Instructional Leadership Training. Respondents, particularly teachers, indicated that appraisers need additional training in recognizing effective practices, providing feedback about effective teaching practices, and providing suggestions for alternate teaching strategies;

Application of TTAS to Diverse Set of Instructional Models and to Higher Order Thinking Skills. There was strong consensus among respondents that TTAS seemed applicable to only one model of teaching (i.e., direct instruction) and to basic thinking skills;

Data Collection and Analysis. Respondents indicated that appraisers need additional training on collecting cumulative data about a teacher's teaching performance; andPre and Post Observation Conferences. Respondents and advisory group members indicated that appraisers need additional training on conducting effective pre and post observation conferences.

BEST COPY AVAILABLE

The following two year action plan has been developed and is being implemented by the MAPD staff to address the following areas of improvement:

1990

(1) Develop training materials to help teachers and appraisers better understand performance indicators with a high level of subjectivity.

(2) Develop training materials to help teachers and appraisers better understand the concept of exceptional quality in teaching.

(3) Develop new videotape proficiencies and written test.

(4) Deliver 1990 Recertification training to appraisers and teachers using the above new materials.

(5) Modify TTAS Certification training materials to incorporate 1990 Recertification training and to provide more comprehensive and substantive Trainer Notes to trainers.

(6) Assess testing practices and inter-rater reliability activities in the TTAS Certification training program.

(7) Merge TTAS Appraiser's Manual and Teacher's TTAS Manual into one Teacher/Appraiser TTAS Manual so that appraisers and teachers have the same information about TTAS practices and procedures.

(Note: All seven projects scheduled for 1990 were successfully completed.)

BEST COPY AVAILABLE

1991

The following improvements are targeted for completion in 1991:

(1) Develop and add videotape inter-rater reliability activities to TTAS Certification training program.

(2) Conduct survey to determine effectiveness of Field Practice activities in school districts (to increase inter-rater reliability).

(3) Conduct survey to determine how to improve delivery of training programs to teachers.

(4) Develop six major training modules to provide continuous, in-depth training to teachers and appraisers, as follows:

  a. TTAS training for teachers;
  b. Data collection and analysis;
  c. Application of TTAS to diverse set of teaching models;
  d. Conference skills;
  e. Mentoring and peer coaching; and
  f. Application of TTAS to higher order thinking skills.

(5) Update research bases of TTAS.

In summary, the TTAS staff is committed to helping teachers and appraisers through in-depth and continuous training to develop a deeper understanding of, and a more collegial approach, to the TTAS. Moreover, there is commitment to help teachers and appraisers apply TTAS to a diverse set of teaching models that include higher order thinking skills.

APPENDIX G

Independent Evaluation of the Texas Teacher Appraisal System

INDEPENDENT EVALUATION OF THE TEXAS TEACHER APPRAISAL SYSTEM

INTRODUCTION

The Texas Education Agency (TEA) selected MGT of America., Inc., to conduct an evaluation of the Texas Teacher Appraisal System (TTAS). The study was conducted in cooperation with the TEA with MGT performing the following major activities:

- Identifying the major issues to consider in the evaluation through interviews with TEA officials and representatives of superintendent, principal, and teacher organizations
- Developing an evaluation design in cooperation with TEA to assess the issues identified
- Designing interview and survey instruments to be used with school district administrators and teachers
- Conducting interviews with administrators and teachers (TEA conducted the surveys of school district administrators and teachers)
- Analyzing qualitative and quantitative data (TEA also analyzed some quantitative data)
- Conducting a national review of research and practices on performance appraisals in both the public and private sectors
- Preparing an independent evaluation of the TTAS.

The TEA organized the findings from these activities into a draft final report with recommendations. MGT utilized the data and information collected in the above activities to prepare this independent evaluation of the TTAS. MGT also utilized the services of two nationally recognized experts in teacher performance evaluation systems, Dr. Richard Brandt and Dr. Russell French, to assist in the preparation of this independent evaluation.

Prior to preparing their respective parts of the final report, the TEA staff, MGT consultants, and the two expert advisors met and jointly reviewed the patterns of responses from the various data sources. The issues which emerged from this analysis were organized under the following topical themes:

- TTAS as a system

- Implementation of the TTAS
- Subjectivity or bias in the TTAS
- Uses of the TTAS
- Impact of the TTAS

To facilitate the preparation of the final report, both the TEA and MGT agreed to organize their respective sections of the final report according to these six themes. MGT's independent evaluation of the TTAS is summarized in the following sections.

THE TTAS AS A SYSTEM

In 1984, the Texas Legislature directed the State Board of Education to adopt an appraisal process and criteria to be used in the appraisal of teachers for career ladder purposes. In addition, the appraisals were to be used for the improvement of instruction through staff development, and for contract renewal purposes. The TEA, after extensive review of the research on effective teaching and teacher evaluation systems in other states, developed and implemented the TTAS in response to the directives of the Legislature and the State Board of Education.

The development and implementation of the TTAS by the TEA incorporated the essential elements of an up-to-date performance appraisal system for teachers:

- The knowledge and skills to be evaluated were based on research on effective teaching and through consensus agreement of persons to be evaluated
- The observation and evaluation instruments were developed, piloted, revised, and validated to measure the required knowledge and skills
- A standardized scoring process was developed and validated to translate measurement data into evaluation decisions
- Evaluator training materials and procedures were developed and trainers and evaluators were trained
- The evaluation system was field tested and reliability studies were conducted

- Instruments and procedures were modified and standards set as a result of the field tests
- Orientation materials were developed and orientations conducted for persons to be evaluated
- Evaluations were conducted and changes made to the system based on this experience
- Training updates were consistently conducted for all trainers and evaluators.

We find that the TTAS as a performance appraisal system is basically sound in overall design and its continued development and use should be encouraged. We do not offer any recommendations on the TTAS as a system, but we do offer some recommendations on various components of the system in the sections which follow.

IMPLEMENTATION OF THE TTAS

Implementation of the TTAS in the 1,058 school districts in Texas was a monumental undertaking and the TEA is commended for its accomplishment. The creditability of the TTAS and its implementation has been clouded by the concurrent implementation of the career ladder program. Creditability has also suffered because of the state mandate that all observations must be announced to the teacher beforehand. Our major concerns with the implementation focus on the frequency and thoroughness of the appraisals, the training of evaluators, and the feedback provided to teachers.

The frequency of appraisals is related to the purpose for which the evaluation is conducted. Currently each teacher must be appraised annually in the performance of his or her duties. Appraisals are required not fewer than:

- two times each year for probationary teachers and for teachers on level one of the career ladder
- once each year for teachers on levels two, three, and four of the career ladder whose performance, on the most recent appraisal, was evaluated as exceeding expectations or clearly outstanding.

The thoroughness of appraisals is dependent both on the number of appraisals and the number of appraisers. Currently the process requires at least two appraisers for each appraisal period. One of the appraisers must be the teacher's supervisor (the principal) and the other may be another administrator or teacher who has received instructional leadership training and the uniform appraiser training. Teachers used as appraisers must have a career ladder assignment at least as high as the teacher to be appraised. Teachers cannot appraise other teachers on the same campus of their teaching assignment.

Principals, especially those with large faculties, must spend an inordinate amount of time conducting teacher appraisals. Yet, making important decisions concerning employment, staff development, and career ladder status based on just two classroom observations is questionable. School districts may assign more appraisers and conduct more than two appraisals per teacher. However, we found this to be a rare exception in practice.

The accuracy of current appraisals is questioned when we consider that out of almost 180,000 teachers who have been appraised with the TTAS, less than one percent were identified as unsatisfactory or performing below expectations and a very large percentage were rated as clearly outstanding.

The primary purpose for the TTAS is to appraise the performance of teachers for career ladder purposes. A secondary purpose is to use information from the appraisals for improvement of instruction. This implies that specific feedback will be provided to the teacher following each appraisal and that staff development programs will be available to instruct the teachers in the specific skills that are assessed by the appraisal. Several citations in the Texas Education Code and State Board of Education Rules on appraisal of certified personnel address this issue, but give the impression that feedback and staff development is intended only for the seriously deficient teacher:

- **Texas Education Code 13.302(e).** The instrument -- be used to assess specific skills primarily for the purpose of remediation and improvement.

- **Rule 149.41(b).** Results -- shall be used for career ladder and staff development purposes.

- **Rule 149.43(c)(9).** If such documentation would influence the teachers appraisal, the documentation must be shared in writing with the teacher within seven working days of the occurrence --

- **Rule 149(c)(10).** Following each formal observation, an appraiser must conduct a post-observation conference -- if the performance is less than meets expectations -- regardless of the performance, each teacher supervisor must conduct a post-observation conference -- at the conclusion of the first appraisal period, a conference will be held at the request of either the teacher or the appraiser.

In both our surveys and interviews, superintendents, principals, and teachers valued post-conferences as the most valuable feature of the TTAS and all three groups gave high marks for the TTAS's ability to assess basic teaching practices for most subjects and teaching at different grade levels. A majority of all respondents also agreed that the TTAS could assess teaching of critical thinking skills.

Strong evidence exists that post-observation conferences for all teachers and readily available staff development programs to address the essential knowledge and skills covered in the TTAS should not be viewed as a secondary purpose for TTAS.

A number of states require that teachers be evaluated at least every three years. This permits the principal to focus annual evaluations on a manageable number of teachers who may require appraisals during the year (i.e., teachers on probation and those seeking higher career ladder status).

We offer the following recommendations based on the findings and observations stated above.

Recommendation 7.1: **Remove the requirement that all teachers be evaluated annually. Those evaluated annually should include probationary teachers and teachers seeking higher career ladder levels.**

Recommendation 7.2: **Increase the number of appraisals for teachers on probation and those seeking higher career ladder status to three. They should be conducted by three different appraisers, one of whom is the principal.**

Our analysis of scores indicated that the TTAS was not providing discriminating scores at the middle to lower end of the scale. In fact, most scores were in the "exceeds expectations" and "clearly outstanding" levels. Observations and comments indicated that the principal was under undue pressure because of the career ladder to inflate scores. This following recommendation should alleviate some of this pressure and yield more reliable scores.

Recommendation 7.3: **Revise the statutes and rules to reflect that the primary purpose of the TTAS is to improve instruction, and its secondary purposes are for contract renewal and career ladder placement.**

The TTAS was initiated as a part of the career ladder legislation, but superintendents, principals and teachers now value it more for its potential to improve instruction.

SUBJECTIVITY OR BIAS IN THE TTAS

Most subjectivity or bias is related to implementation of the TTAS rather than to the instrument. Thus, our recommendations above are also applicable in this section.

Most persons interviewed (89%) believe that TTAS scores are influenced by factors other than teaching performance. Factors cited were:

- subjectivity of the TTAS system
- limited career ladder funds
- appraisers lack of knowledge of subject matter being taught.

A majority of the survey respondents (72%) believe appraisers are capable of using the TTAS accurately as it is presently designed. In another series of survey questions, a majority of both administrators and teachers (63%) believe there are problems with either the TTAS or its implementation.

The recent change requiring that all classroom observations be announced prior to the visitations diminished the reliability of the TTAS according to a number of the administrators and some teachers.

Questions on the reliability of performance appraisal systems are to be expected, especially a developing system being implemented in 1,058 different school districts. We believe the only way to effectively address these questions is to periodically conduct independent reliability checks on the system.

Recommendation 7.4: **Design and conduct a reliability study of the TTAS in several school districts in Texas matching the data collected by trained independent appraisers during unscheduled classroom observations with data collected by district appraisers operating under current procedures.**

USES OF TTAS

As indicated earlier, the TTAS originated as a system to implement the career ladder with secondary objectives related to employment and staff development. We discussed and made a recommendation above on the use of the TTAS for staff development. Our interviews indicated a great majority (78%) of district personnel believed the TTAS should be linked to contract renewal or termination. Barely a majority (51%) of all survey respondents, and only 37 percent of the teachers would link the TTAS to contracts or dismissal.

We attempted, without success, to evaluate the TTAS apart from the career ladder. The two are closely integrated in the minds of school district personnel. Survey respondents were split in their opinions of linking the TTAS to the career ladder. A narrow majority (51%) of the survey respondents rejected the TTAS as appropriate for appraisal of master teachers.

The lack of discriminating power of the TTAS is a serious deficiency. More extensive data are need. Other states, Tennessee for example, use multiple data sources for the career ladder which become even more important as the candidates approach the top level of the ladder (Master Teacher). The TTAS included another data source (Teacher Assessment of Instructional Goals and Outcomes) but it was discontinued as not contributing meaningful data.

Exceptional quality (EQ) points, awarded at the domain level, are another data source in current use. A number of persons interviewed at the state and district levels indicated that EQ points could make the TTAS more discriminating if they were awarded at the indicator, rather that domain level. The greatest objection to awarding EQ points at the indicator level is the increase in documentation required. If the number of required annual appraisals is reduced (recommendation 7.1), the objection to awarding EQ points at the indicator level would be reduced. Our recommendation on EQ points is as follows:

Recommendation 7.5: **Award Exceptional Quality points at the indicator level with appraisers providing evidence supporting the award. Training for appraisers on EQ standards, procedures, delineation of behavior patterns required for awards, and a structured process for the different appraisers to reach consensus on the awards to an individual should be provided.**

A written examination is proposed as another data source for the master teacher level in Texas. The validity and reliability of a written test for this purpose is yet to be assessed.

The TTAS should be continued as one of several data sources for the career ladder. If the new Texas Academic Excellence Indicator System is a valid measure of student higher order performance, then the results of this measure over time for students of a teacher aspiring to advance on the career ladder should be considered as another data source.

In addition to multiple data sources, some additional attention needs to be given to the expertise of appraisers when the TTAS is used for career ladder purposes. Certainly the expertise of the appraiser should not be questionable, especially at higher levels of the career ladder. Recent research (Berliner, 1988) indicates major differences in ratings of teacher effectiveness associated with the different levels of pedagogical expertise of the appraisers. Some TTAS appraisers now have several years of experience with the system. Some of these are probably much more accurate in their appraisal than others. It should be possible to identify a cadre of the most accurate appraisers and then use these individuals to assess teachers at the higher levels of the career ladder. When master teachers have been named they should be utilized as appraisers for the master teacher level.

We offered some recommendations earlier that relate to the use of the TTAS. Additional recommendations follow:

Recommendation 7.6: **The TTAS should be considered a developing system with additional data sources added to the system as they are developed and validated. Consideration should be given to inclusion of data regarding innovative teaching practices, teacher leadership and other performance dimensions corresponding to current educational concerns and the most recent research on teaching.**

Recommendation 7.7: **The TTAS should continue to be used for the purposes of instructional improvement/staff development decisions, employment decisions, and career ladder decisions.**

Recommendation 7.8: **A cadre of expert appraisers should be identified/trained and utilized for appraisals at higher levels of the career ladder. Consideration should be given for one or two members of the career ladder appraisal team to be peer teachers from outside the candidates' own district. The influence of the principal in career ladder appraisals should be reduced.**

IMPACT OF THE TTAS

Teachers and administrators reported that the TTAS has created a commonality of language and effective teaching practices in Texas. Principals reported that it has significantly increased the amount of time they spend in the classroom. A majority of the teachers and administrators in both interviews and surveys reported that the TTAS had:

- improved classroom teaching methods
- identified ineffective teaching practices
- identified exceptional quality of teaching practices

The TTAS has systemized the performance appraisal process and demonstrated transportability of performance appraisals across the 1,058 school districts in Texas. A majority of the administrators and 43 percent of the teachers in the survey reported that the TTAS resulted in more fairness in the teacher appraisal than earlier evaluation systems.

The reporting of good feelings by teachers and administrators that the TTAS has improved instruction is not sufficient to satisfy the public and policymakers that the significant amount of funding required to further develop the TTAS and career ladder is cost-effective. Some reliable outcome measures are needed to support the "good feelings".

Student performance on the Texas Educational Assessment of Minimum Skills (TEAMS) was analyzed in relation to teacher performance on the TTAS. While some gains in TEAMS scores are evident, it would be speculative to attribute these gains to improved teaching as a result of the TTAS.

A more promising approach than trying to relate aggregated student performance data to the TTAS is to relate school performance data to the performance of teachers. A school that has a significant number of faculty scoring at the higher levels on the TTAS should be able to demonstrate some correlation to increased student performance. The recent introduction of the new Texas Assessment of Academic Skills (TAAS) elevates state student testing above minimum levels. This should enhance the opportunity for some meaningful experimentation with higher levels of performance on the TTAS.

Recommendation 7.9: **Design and fund experimental school-based incentive programs which relate teacher incentives to student outcomes and student outcomes to teacher performance.**

The Texas Education Agency currently publishes average TTAS scores by school district in Snapshot: School District Profiles. No explanation is given for the variation in mean TTAS scores from district to district. This reporting would probably have greater impact if the mean and range of TTAS scores were reported by district.

This reporting concept applied to TTAS appraisers could help reduce the current TTAS score inflation. If an appraiser had access to the mean, median, and range of TTAS scores of other appraises, it would permit a self-assessment of the appraisers performance. It would also permit the TEA to target appraisers who might benefit from additional appraiser training.

Recommendation 7.10: **Consideration should be given to a statewide system to monitor the TTAS scores awarded by all appraisers with copies of the mean, median and range of TTAS scores of all appraisers provided to all appraisers and school district superintendents.**

COMPLIANCE STATEMENT

TITLE VI, CIVIL RIGHTS ACT OF 1964; THE MODIFIED COURT ORDER, CIVIL ACTION 5281, FEDERAL DISTRICT COURT, EASTERN DISTRICT OF TEXAS, TYLER DIVISION

Reviews of local education agencies pertaining to compliance with Title VI Civil Rights Act of 1964 and with specific requirements of the Modified Court Order, Civil Action No. 5281, Federal District Court, Eastern District of Texas, Tyler Division are conducted periodically by staff representatives of the Texas Education Agency. These reviews cover at least the following policies and practices:

(1) acceptance policies on student transfers from other school districts;

(2) operation of school bus routes or runs on a non-segregated basis;

(3) nondiscrimination in extracurricular activities and the use of school facilities;

(4) nondiscriminatory practices in the hiring, assigning, promoting, paying, demoting, reassigning, or dismissing of faculty and staff members who work with children;

(5) enrollment and assignment of students without discrimination on the basis of race, color, or national origin;

(6) nondiscriminatory practices relating to the use of a student's first language; and

(7) evidence of published procedures for hearing complaints and grievances.

In addition to conducting reviews, the Texas Education Agency staff representatives check complaints of discrimination made by a citizen or citizens residing in a school district where it is alleged discriminatory practices have occurred or are occurring.

Where a violation of Title VI of the Civil Rights Act is found, the findings are reported to the Office for Civil Rights, U.S. Department of Education.

If there is a direct violation of the Court Order in Civil Action No. 5281 that cannot be cleared through negotiation, the sanctions required by the Court Order are applied.

TITLE VII, CIVIL RIGHTS ACT OF 1964; EXECUTIVE ORDERS 11246 AND 11375; TITLE IX, 1973 EDUCATION AMENDMENTS; REHABILITATION ACT OF 1973 AS AMENDED; 1974 AMENDMENTS TO THE WAGE-HOUR LAW EXPANDING THE AGE DISCRIMINATION IN EMPLOYMENT ACT OF 1967; AND VIETNAM ERA VETERANS READJUSTMENT ASSISTANCE ACT OF 1972 AS AMENDED IN 1974.

It is the policy of the Texas Education Agency to comply fully with the nondiscrimination provisions of all federal and state laws and regulations by assuring that no person shall be excluded from consideration for recruitment, selection, appointment, training, promotion, retention, or any other personnel action, or be denied any benefits or participation in any programs or activities which it operates on the grounds of race, religion, color, national origin, sex, handicap, age, or veteran status (except where age, sex, or handicap constitute a bona fide occupational qualification necessary to proper and efficient administration). The Texas Education Agency makes positive efforts to employ and advance in employment all protected groups.

TEXAS EDUCATION AGENCY
1701 NORTH CONGRESS AVENUE
AUSTIN, TEXAS 78701-1494

GE1 543 03

BEST COPY AVAILABLE